

B9

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
28 August 2003 (28.08.2003)

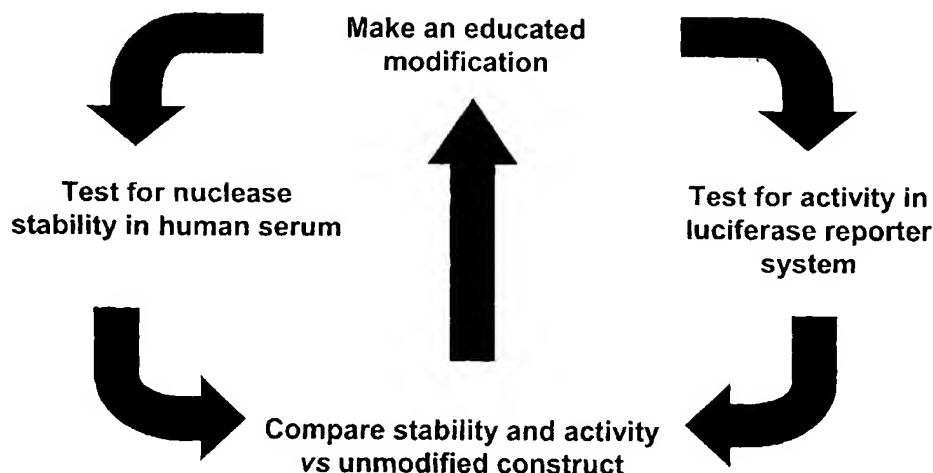
PCT

(10) International Publication Number  
**WO 03/070193 A2**

(51) International Patent Classification <sup>7</sup> :	A61K	60/440,129	15 January 2003 (15.01.2003)	US
(21) International Application Number:	PCT/US03/05190	(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:		
(22) International Filing Date:	20 February 2003 (20.02.2003)	US	60/398,036 (CON)	
(25) Filing Language:	English	Filed on	23 July 2002 (23.07.2002)	
(26) Publication Language:	English	US	10/225,023 (CON)	
(30) Priority Data:		Filed on	21 August 2002 (21.08.2002)	
60/358,580	20 February 2002 (20.02.2002)	US	10/157,580 (CON)	
60/363,124	11 March 2002 (11.03.2002)	US	29 May 2002 (29.05.2002)	
60/374,722	22 April 2002 (22.04.2002)	US	US	60/294,140 (CON)
10/157,580	29 May 2002 (29.05.2002)	US	Filed on	29 May 2001 (29.05.2001)
60/386,782	6 June 2002 (06.06.2002)	US	US	60/374,722 (CON)
60/398,036	23 July 2002 (23.07.2002)	US	Filed on	22 April 2002 (22.04.2002)
10/225,023	21 August 2002 (21.08.2002)	US	US	60/358,580 (CON)
60/406,784	29 August 2002 (29.08.2002)	US	Filed on	20 February 2002 (20.02.2002)
60/408,378	5 September 2002 (05.09.2002)	US	US	60/363,124 (CON)
60/409,293	9 September 2002 (09.09.2002)	US	Filed on	11 March 2002 (11.03.2002)
				60/386,782 (CON)
				6 June 2002 (06.06.2002)
				60/406,784 (CON)
				29 August 2002 (29.08.2002)

[Continued on next page]

(54) Title: RNA INTERFERENCE MEDIATED INHIBITION OF HIV GENE EXPRESSION USING SHORT INTERFERING NUCLEIC ACID (siNA)



(57) Abstract: The present invention concerns methods and reagents useful in modulating human immunodeficiency virus (HIV) gene expression in a variety of applications, including use in therapeutic, diagnostic, target validation, and genomic discovery applications. Specifically, the invention relates to small nucleic acid molecules, such as short interfering nucleic acid (siNA), short interfering RNA (siRNA), double-stranded RNA (dsRNA), micro-RNA (miRNA), and short hairpin RNA (shRNA) molecules capable of mediating RNA interference (RNAi) against HIV gene expression and/or activity. The small nucleic acid molecules are useful in the treatment and diagnosis of HIV infection, AIDS, and any other disease or condition that responds to modulation of HIV expression or activity.



WO 03/070193 A2



US 60/408,378 (CON)  
Filed on 5 September 2002 (05.09.2002)  
US 60/409,293 (CON)  
Filed on 9 September 2002 (09.09.2002)  
US 60/440,129 (CON)  
Filed on 15 January 2003 (15.01.2003)

(71) **Applicant** (*for all designated States except US*): **SIRNA THERAPEUTICS, INC** [US/US]; 2950 Wildemess Place, Boulder, CO 80301 (US).

(71) **Applicant and**

(72) **Inventor** (*for US only*): **MCSWIGGEN, James** [US/US]; 4866 Franklin Drive, Boulder, CO 80301 (US).

(72) **Inventors; and**

(75) **Inventors/Applicants** (*for US only*): **BEIGELMAN, Leonid** [US/US]; 5530 Colt Drive, Longmont, CO 80503 (US). **MACEJAK, Dennis** [US/US]; 6595 Union Street, Arvada, CO 80004 (US).

(74) **Agent**: **TERPSTRA, Anita, J.**; McDonnell Boehnen Hulbert & Berghoff, 300 South Wacker Drive, Suite 3200, Chicago, IL 60606 (US).

(81) **Designated States** (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) **Designated States** (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

**Published:**

— *without international search report and to be republished upon receipt of that report*

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

## RNA INTERFERENCE MEDIATED INHIBITION OF HIV GENE EXPRESSION USING SHORT INTERFERING NUCLEIC ACID (siNA)

This invention claims the benefit of McSwiggen, USSN 60/398,036 filed July 23, 2002, of McSwiggen, USSN 10/225,023 filed August 21, 2002, of McSwiggen, USSN 5 10/157,580 filed May 29, 2002, which claims the benefit of McSwiggen USSN 60/294,140 filed May 29, 2001, of McSwiggen USSN 60/374,722 filed April 22, 2002, of Beigelman USSN 60/358,580 filed February 20, 2002, of Beigelman USSN 60/363,124 filed March 11, 2002, of Beigelman USSN 60/386,782 filed June 6, 2002, of Beigelman USSN 60/406,784 filed August 29, 2002, of Beigelman USSN 60/408,378 10 filed September 5, 2002, of Beigelman USSN 60/409,293 filed September 9, 2002, and of Beigelman USSN 60/440,129 filed January 15, 2003. These applications are hereby incorporated by reference herein in their entireties, including the drawings.

### Field Of The Invention

The present invention concerns compounds, compositions, and methods for the 15 study, diagnosis, and treatment of conditions and diseases that respond to the modulation of human immunodeficiency virus (HIV) gene expression and/or activity. The present invention also concerns compounds, compositions, and methods relating to conditions and diseases that respond to the modulation of expression and/or activity of genes involved in human immunodeficiency virus (HIV) expression and activity pathways. 20 Specifically, the invention relates to small nucleic acid molecules, such as short interfering nucleic acid (siNA), short interfering RNA (siRNA), double-stranded RNA (dsRNA), micro-RNA (miRNA), and short hairpin RNA (shRNA) molecules capable of mediating RNA interference (RNAi) against HIV gene expression.

### Background Of The Invention

25 The following is a discussion of relevant art pertaining to RNAi. The discussion is provided only for understanding of the invention that follows. The summary is not an admission that any of the work described below is prior art to the claimed invention.

RNA interference refers to the process of sequence-specific post-transcriptional gene silencing in animals mediated by short interfering RNAs (siRNAs) (Fire *et al.*, 30 1998, *Nature*, 391, 806). The corresponding process in plants is commonly referred to as

post-transcriptional gene silencing or RNA silencing and is also referred to as quelling in fungi. The process of post-transcriptional gene silencing is thought to be an evolutionarily-conserved cellular defense mechanism used to prevent the expression of foreign genes and is commonly shared by diverse flora and phyla (Fire *et al.*, 1999, *Trends Genet.*, 15, 358). Such protection from foreign gene expression may have evolved in response to the production of double-stranded RNAs (dsRNAs) derived from viral infection or from the random integration of transposon elements into a host genome via a cellular response that specifically destroys homologous single-stranded RNA or viral genomic RNA. The presence of dsRNA in cells triggers the RNAi response though a mechanism that has yet to be fully characterized. This mechanism appears to be different from the interferon response that results from dsRNA-mediated activation of protein kinase PKR and 2',5'-oligoadenylate synthetase resulting in non-specific cleavage of mRNA by ribonuclease L.

The presence of long dsRNAs in cells stimulates the activity of a ribonuclease III enzyme referred to as dicer. Dicer is involved in the processing of the dsRNA into short pieces of dsRNA known as short interfering RNAs (siRNAs) (Berstein *et al.*, 2001, *Nature*, 409, 363). Short interfering RNAs derived from dicer activity are typically about 21 to about 23 nucleotides in length and comprise about 19 base pair duplexes (Elbashir *et al.*, 2001, *Genes Dev.*, 15, 188). Dicer has also been implicated in the excision of 21- and 22-nucleotide small temporal RNAs (stRNAs) from precursor RNA of conserved structure that are implicated in translational control (Hutvagner *et al.*, 2001, *Science*, 293, 834). The RNAi response also features an endonuclease complex, commonly referred to as an RNA-induced silencing complex (RISC), which mediates cleavage of single-stranded RNA having sequence complementary to the antisense strand of the siRNA duplex. Cleavage of the target RNA takes place in the middle of the region complementary to the antisense strand of the siRNA duplex (Elbashir *et al.*, 2001, *Genes Dev.*, 15, 188).

RNAi has been studied in a variety of systems. Fire *et al.*, 1998, *Nature*, 391, 806, were the first to observe RNAi in *C. elegans*. Wianny and Goetz, 1999, *Nature Cell Biol.*, 2, 70, describe RNAi mediated by dsRNA in mouse embryos. Hammond *et al.*, 2000, *Nature*, 404, 293, describe RNAi in *Drosophila* cells transfected with dsRNA. Elbashir *et al.*, 2001, *Nature*, 411, 494, describe RNAi induced by introduction of



duplexes of synthetic 21-nucleotide RNAs in cultured mammalian cells including human embryonic kidney and HeLa cells. Recent work in *Drosophila* embryonic lysates (Elbashir *et al.*, 2001, *EMBO J.*, 20, 6877) has revealed certain requirements for siRNA length, structure, chemical composition, and sequence that are essential to mediate efficient RNAi activity. These studies have shown that 21-nucleotide siRNA duplexes are most active when containing 3'-terminal dinucleotide overhangs. Furthermore, complete substitution of one or both siRNA strands with 2'-deoxy (2'-H) or 2'-O-methyl nucleotides abolishes RNAi activity, whereas substitution of the 3'-terminal siRNA overhang nucleotides with 2'-deoxy nucleotides (2'-H) was shown to be tolerated. Single mismatch sequences in the center of the siRNA duplex were also shown to abolish RNAi activity. In addition, these studies also indicate that the position of the cleavage site in the target RNA is defined by the 5'-end of the siRNA guide sequence rather than the 3'-end of the guide sequence (Elbashir *et al.*, 2001, *EMBO J.*, 20, 6877). Other studies have indicated that a 5'-phosphate on the target-complementary strand of a siRNA duplex is required for siRNA activity and that ATP is utilized to maintain the 5'-phosphate moiety on the siRNA (Nykanen *et al.*, 2001, *Cell*, 107, 309).

Studies have shown that replacing the 3'-terminal nucleotide overhanging segments of a 21-mer siRNA duplex having two -nucleotide 3'-overhangs with deoxyribonucleotides does not have an adverse effect on RNAi activity. Replacing up to four nucleotides on each end of the siRNA with deoxyribonucleotides has been reported to be well tolerated, whereas complete substitution with deoxyribonucleotides results in no RNAi activity (Elbashir *et al.*, 2001, *EMBO J.*, 20, 6877). In addition, Elbashir *et al.*, *supra*, also report that substitution of siRNA with 2'-O-methyl nucleotides completely abolishes RNAi activity. Li *et al.*, International PCT Publication No. WO 00/44914, and Beach *et al.*, International PCT Publication No. WO 01/68836 preliminarily suggest that siRNA may include modifications to either the phosphate-sugar backbone or the nucleoside to include at least one of a nitrogen or sulfur heteroatom, however, neither application postulates to what extent such modifications would be tolerated in siRNA molecules, nor provides any further guidance or examples of such modified siRNA. Kreutzer *et al.*, Canadian Patent Application No. 2,359,180, also describe certain chemical modifications for use in dsRNA constructs in order to counteract activation of double-stranded RNA-dependent protein kinase PKR, specifically 2'-amino or 2'-O-methyl nucleotides, and nucleotides containing a 2'-O or 4'-C methylene bridge.

However, Kreutzer *et al.* similarly fails to provide examples or guidance as to what extent these modifications would be tolerated in siRNA molecules.

Parrish *et al.*, 2000, *Molecular Cell*, 6, 1977-1087, tested certain chemical modifications targeting the *unc-22* gene in *C. elegans* using long (>25 nt) siRNA transcripts. The authors describe the introduction of thiophosphate residues into these  
5 siRNA transcripts by incorporating thiophosphate nucleotide analogs with T7 and T3 RNA polymerase and observed that RNAs with two phosphorothioate modified bases also had substantial decreases in effectiveness as RNAi. Further, Parrish *et al.* reported that phosphorothioate modification of more than two residues greatly destabilized the  
10 RNAs *in vitro* such that interference activities could not be assayed. *Id.* at 1081. The authors also tested certain modifications at the 2'-position of the nucleotide sugar in the long siRNA transcripts and found that substituting deoxynucleotides for ribonucleotides produced a substantial decrease in interference activity, especially in the case of Uridine to Thymidine and/or Cytidine to deoxy-Cytidine substitutions. *Id.* In addition, the  
15 authors tested certain base modifications, including substituting, in sense and antisense strands of the siRNA, 4-thiouracil, 5-bromouracil, 5-iodouracil, and 3-(aminoallyl)uracil for uracil, and inosine for guanosine. Whereas 4-thiouracil and 5-bromouracil substitution appeared to be tolerated, Parrish reported that inosine produced a substantial decrease in interference activity when incorporated in either strand. Parrish also reported  
20 that incorporation of 5-iodouracil and 3-(aminoallyl)uracil in the antisense strand resulted in a substantial decrease in RNAi activity as well.

The use of longer dsRNA has been described. For example, Beach *et al.*, International PCT Publication No. WO 01/68836, describes specific methods for attenuating gene expression using endogenously-derived dsRNA. Tuschl *et al.*,  
25 International PCT Publication No. WO 01/75164, describe a *Drosophila in vitro* RNAi system and the use of specific siRNA molecules for certain functional genomic and certain therapeutic applications; although Tuschl, 2001, *Chem. Biochem.*, 2, 239-245, doubts that RNAi can be used to cure genetic diseases or viral infection due to the danger of activating interferon response. Li *et al.*, International PCT Publication No. WO  
30 00/44914, describe the use of specific dsRNAs for attenuating the expression of certain target genes. Zernicka-Goetz *et al.*, International PCT Publication No. WO 01/36646, describe certain methods for inhibiting the expression of particular genes in mammalian

cells using certain dsRNA molecules. Fire *et al.*, International PCT Publication No. WO 99/32619, describe particular methods for introducing certain dsRNA molecules into cells for use in inhibiting gene expression. Plaetinck *et al.*, International PCT Publication No. WO 00/01846, describe certain methods for identifying specific genes  
5 responsible for conferring a particular phenotype in a cell using specific dsRNA molecules. Mello *et al.*, International PCT Publication No. WO 01/29058, describe the identification of specific genes involved in dsRNA-mediated RNAi. Deschamps Depaillette *et al.*, International PCT Publication No. WO 99/07409, describe specific compositions consisting of particular dsRNA molecules combined with certain anti-viral  
10 agents. Waterhouse *et al.*, International PCT Publication No. 99/53050, describe certain methods for decreasing the phenotypic expression of a nucleic acid in plant cells using certain dsRNAs. Driscoll *et al.*, International PCT Publication No. WO 01/49844, describe specific DNA constructs for use in facilitating gene silencing in targeted organisms.

15 Others have reported on various RNAi and gene-silencing systems. For example, Parrish *et al.*, 2000, *Molecular Cell*, 6, 1977-1087, describe specific chemically-modified siRNA constructs targeting the unc-22 gene of *C. elegans*. Grossniklaus, International PCT Publication No. WO 01/38551, describes certain methods for regulating polycomb gene expression in plants using certain dsRNAs. Churikov *et al.*, International PCT  
20 Publication No. WO 01/42443, describe certain methods for modifying genetic characteristics of an organism using certain dsRNAs. Cogoni *et al.*, International PCT Publication No. WO 01/53475, describe certain methods for isolating a Neurospora silencing gene and uses thereof. Reed *et al.*, International PCT Publication No. WO 01/68836, describe certain methods for gene silencing in plants. Honer *et al.*,  
25 International PCT Publication No. WO 01/70944, describe certain methods of drug screening using transgenic nematodes as Parkinson's Disease models using certain dsRNAs. Deak *et al.*, International PCT Publication No. WO 01/72774, describe certain *Drosophila*-derived gene products that may be related to RNAi in *Drosophila*. Arndt *et al.*, International PCT Publication No. WO 01/92513 describe certain methods for  
30 mediating gene suppression by using factors that enhance RNAi. Tuschl *et al.*, International PCT Publication No. WO 02/44321, describe certain synthetic siRNA constructs. Pachuk *et al.*, International PCT Publication No. WO 00/63364, and Satishchandran *et al.*, International PCT Publication No. WO 01/04313, describe certain

- methods and compositions for inhibiting the function of certain polynucleotide sequences using certain dsRNAs. Echeverri *et al.*, International PCT Publication No. WO 02/38805, describe certain *C. elegans* genes identified via RNAi. Kreutzer *et al.*, International PCT Publications Nos. WO 02/055692, WO 02/055693, and EP 1144623
- 5 B1 describes certain methods for inhibiting gene expression using RNAi. Graham *et al.*, International PCT Publications Nos. WO 99/49029 and WO 01/70949, and AU 4037501 describe certain vector expressed siRNA molecules. Fire *et al.*, US 6,506,559, describe certain methods for inhibiting gene expression in vitro using certain long dsRNA (greater than 25 nucleotide) constructs that mediate RNAi.
- 10 Acquired immunodeficiency syndrome (AIDS) is thought to be caused by infection with the human immunodeficiency virus, for example, HIV-1. Draper *et al.*, U.S. Patent Nos. 6,159,692, 5,972,704, 5,693,535, and International PCT Publication Nos. WO 93/23569 and WO 95/04818, describes enzymatic nucleic acid molecules targeting HIV. Novina *et al.*, 2002, *Nature Medicine*, 8, 681-686, describes certain siRNA constructs
- 15 targeting HIV-1 infection. Lee *et al.*, 2002, *Nature Biotechnology*, 19, 500-505, describes certain siRNA targeted against HIV-1 rev.

## SUMMARY OF THE INVENTION

- This invention relates to compounds, compositions, and methods useful for modulating the expression of genes, such as those genes associated with the development
- 20 or maintenance of human immunodeficiency virus (HIV) infection and acquired immunodeficiency syndrome (AIDS), by RNA interference (RNAi), using short interfering nucleic acid (siNA) molecules. This invention also relates to compounds, compositions, and methods useful for modulating the expression and activity of HIV genes, or genes involved in HIV pathways or genes involved in HIV activity by RNA
- 25 interference (RNAi) using small nucleic acid molecules, such as short interfering nucleic acid (siNA), short interfering RNA (siRNA), double-stranded RNA (dsRNA), micro-RNA (miRNA), and short hairpin RNA (shRNA) molecules. In particular, the instant invention features small nucleic acid molecules, such as short interfering nucleic acid (siNA), short interfering RNA (siRNA), double-stranded RNA (dsRNA), micro-RNA
- 30 (miRNA), and short hairpin RNA (shRNA) molecules and methods used to modulate the expression of HIV genes. A siNA of the invention can be unmodified or chemically-

modified. A siNA of the instant invention can be chemically synthesized, expressed from a vector or enzymatically synthesized. The instant invention also features various chemically-modified synthetic short interfering nucleic acid (siNA) molecules capable of modulating HIV gene expression or activity in cells by RNA interference (RNAi). The use of chemically-modified siNA improves various properties of native siNA molecules through increased resistance to nuclease degradation *in vivo* and/or through improved cellular uptake. Further, contrary to earlier published studies, siNA having multiple chemical modifications retains its RNAi activity. The siNA molecules of the instant invention provide useful reagents and methods for a variety of therapeutic, diagnostic, target validation, genomic discovery, genetic engineering, and pharmacogenomic applications.

In one embodiment, the invention features a double-stranded short interfering nucleic acid (siNA) molecule that inhibits replication of a human immunodeficiency virus (HIV), wherein one of the strands of the double-stranded siNA molecule is an antisense strand which comprises a nucleotide sequence that is complementary to the nucleotide sequence of an HIV RNA or a portion thereof and the other strand is a sense strand which comprises a nucleotide sequence that is complementary to the nucleotide sequence of the antisense strand. In one embodiment, the HIV RNA comprises HIV-1RNA. In another embodiment, the HIV RNA comprises HIV-2 RNA.

In one embodiment, the invention features a double-stranded short interfering nucleic acid (siNA) molecule that inhibits replication of a human immunodeficiency virus (HIV), wherein one of the strands of the double-stranded siNA molecule is an antisense strand which comprises a nucleotide sequence that is complementary to the nucleotide sequence of an HIV RNA or a portion thereof, and the other strand is a sense strand which comprises a nucleotide sequence that is complementary to the nucleotide sequence of the antisense strand, wherein a majority of the pyrimidine nucleotides present in the double-stranded siNA molecule comprises a sugar modification. In one embodiment, all of the pyrimidine nucleotides present in the double-stranded siNA molecule comprise a sugar modification. In one embodiment, each strand of the double-stranded siNA molecule comprises about 19 to about 29 nucleotides and each strand comprises at least about 19 nucleotides that are complementary to the nucleotides of the other strand. In another embodiment, the double-stranded siNA molecule is assembled

from two oligonucleotide fragments, wherein one fragment comprises nucleotide sequence of the antisense strand of the siNA molecule and the second fragment comprises nucleotide sequence of the sense strand of the siNA molecule. In yet another embodiment, the sense strand of the double-stranded siNA molecule is connected to the antisense strand via a linker molecule, such as a polynucleotide linker or a non-nucleotide linker. In another embodiment, any pyrimidine nucleotides (i.e., one or more or all) present in the sense strand of the double-stranded siNA molecule are 2'-deoxy-2'-fluoro pyrimidine nucleotides and any purine nucleotides (i.e., one or more or all) present in the sense region are 2'-deoxy purine nucleotides. In yet another embodiment, the sense strand of the double-stranded siNA molecule comprises a 3'-end and a 5'-end, wherein a terminal cap moiety (e.g., an inverted deoxy abasic moiety) is present at the 5'-end, the 3'-end, or both of the 5' and 3' ends of the sense strand. In another embodiment, the antisense strand of the double-stranded siNA molecule comprises one or more 2'-deoxy-2'-fluoro pyrimidine nucleotides and one or more 2'-O-methyl purine nucleotides. In yet another embodiment, any pyrimidine nucleotides present in the antisense strand of the double-stranded siNA molecule are 2'-deoxy-2'-fluoro pyrimidine nucleotides and any purine nucleotides present in the antisense strand are 2'-O-methyl purine nucleotides. In another embodiment, the antisense strand of the double-stranded siNA molecule comprises a phosphorothioate internucleotide linkage at the 3' end of the antisense strand. In yet another embodiment, the antisense strand comprises a glyceryl modification at the 3' end of the antisense strand. In still another embodiment, the 5'-end of the antisense strand optionally includes a phosphate group.

In one embodiment, the invention features a double-stranded short interfering nucleic acid (siNA) molecule that inhibits replication of a human immunodeficiency virus (HIV), wherein one of the strands of the double-stranded siNA molecule is an antisense strand which comprises a nucleotide sequence that is complementary to the nucleotide sequence of an HIV RNA or a portion thereof and the other strand is a sense strand which comprises a nucleotide sequence that is complementary to the nucleotide sequence of the antisense strand, wherein a majority of the pyrimidine nucleotides present in the double-stranded siNA molecule comprises a sugar modification, and wherein each of the two strands of said siNA molecule comprises 21 nucleotides. In one embodiment, 21 nucleotides of the antisense strand are base-paired to the nucleotide sequence of the HIV RNA or a portion thereof. In another embodiment, about 19

nucleotides of the antisense strand are base-paired to the nucleotide sequence of the HIV RNA or a portion thereof. In one embodiment, each strand of the siNA molecule is base-paired to the complementary nucleotides of the other strand of the siNA molecule. In another embodiment, about 19 nucleotides of each strand of the siNA molecule are base-paired to the complementary nucleotides of the other strand of the siNA molecule and at least two 3' terminal nucleotides of each strand of the siNA molecule are not base-paired to the nucleotides of the other strand of the siNA molecule. In one embodiment, each of the two 3' terminal nucleotides of each strand of the siNA molecule that are not base-paired are 2'-deoxy-pyrimidines, such as 2'-deoxy-thymidine.

10 In one embodiment, the invention features a double-stranded short interfering nucleic acid (siNA) molecule that inhibits replication of a human immunodeficiency virus (HIV), wherein one of the strands of the double-stranded siNA molecule is an antisense strand which comprises a nucleotide sequence that is complementary to the nucleotide sequence of an HIV RNA or a portion thereof and the other strand is a sense strand which comprises a nucleotide sequence that is complementary to the nucleotide sequence of the antisense strand, wherein a majority of the pyrimidine nucleotides present in the double-stranded siNA molecule comprises a sugar modification and wherein the nucleotide sequence of the antisense strand or a portion thereof is complementary to a nucleotide sequence of the 5'-untranslated region of an HIV RNA or a portion thereof.

25 In another embodiment, the invention features a double-stranded short interfering nucleic acid (siNA) molecule that inhibits replication of a human immunodeficiency virus (HIV), wherein one of the strands of the double-stranded siNA molecule is an antisense strand which comprises a nucleotide sequence that is complementary to the nucleotide sequence of an HIV RNA or a portion thereof, and the other strand is a sense strand which comprises a nucleotide sequence that is complementary to the nucleotide sequence of the antisense strand, wherein a majority of the pyrimidine nucleotides present in the double-stranded siNA molecule comprises a sugar modification and wherein the nucleotide sequence of the antisense strand or a portion thereof is complementary to a nucleotide sequence of an HIV RNA that is present in the RNA of all HIV.

In one embodiment, the invention features a double-stranded short interfering nucleic acid (siNA) molecule that inhibits replication of a human immunodeficiency virus (HIV), wherein one of the strands of the double-stranded siNA molecule is an antisense strand which comprises nucleotide sequence that is complementary to the  
5 nucleotide sequence of an RNA encoding an HIV protein or a fragment thereof and the other strand is a sense strand which comprises a nucleotide sequence that is complementary to the nucleotide sequence of the antisense strand. In one embodiment, a majority of the pyrimidine nucleotides present in the double-stranded siNA molecule comprises a sugar modification.

10 In one embodiment, the nucleotide sequence of the antisense strand or a portion thereof of a siNA molecule of the invention is complementary to the nucleotide sequence of an HIV RNA or a portion thereof that is present in the RNA of all HIV strains.

In one embodiment, the invention features a pharmaceutical composition comprising a siNA molecule of the invention in an acceptable carrier or diluent.

15 In one embodiment, the invention features a medicament comprising an siNA molecule of the invention.

In one embodiment, the invention features an active ingredient comprising an siNA molecule of the invention.

In one embodiment, the invention features the use of a double-stranded short  
20 interfering nucleic acid (siNA) molecule that inhibits replication of a human immunodeficiency virus (HIV), wherein one of the strands of said double-stranded siNA molecule is an antisense strand which comprises a nucleotide sequence that is complementary to the nucleotide sequence of an HIV RNA or a portion thereof and the other strand is a sense strand which comprises a nucleotide sequence that is  
25 complementary to the nucleotide sequence of the antisense strand, wherein a majority of the pyrimidine nucleotides present in said double-stranded siNA molecule comprises a sugar modification.

In one embodiment, the invention features one or more siNA molecules and methods that independently or in combination modulate the expression of gene(s)  
30 encoding HIV and/or HIV polypeptides. Specifically, the present invention features



siNA molecules that modulate the expression of HIV, for example, HIV-1, HIV-2, and related viruses such as FIV-1 and SIV-1; or a specific HIV gene, for example, LTR, *nef*, *vif*, *tat*, or *rev*. In particular embodiments, the invention features nucleic acid-based molecules and methods that modulate the expression of HIV-1 encoded genes, for example Genbank Accession No. AJ302647; HIV-2 gene, for example Genbank Accession No. NC\_001722; FIV-1, for example Genbank Accession No. NC\_001482; SIV-1, for example Genbank Accession No. M66437; *LTR*, for example included in Genbank Accession No. AJ302647; *nef*, for example included in Genbank Accession No. AJ302647; *vif*, for example included in Genbank Accession No. AJ302647; *tat*, for example included in Genbank Accession No. AJ302647; and *rev*, for example included in Genbank Accession No. AJ302647.

In another embodiment, the invention features one or more siNA molecules and methods that independently or in combination modulate the expression of gene(s) encoding the HIV-1 envelope glycoprotein (*env*, for example, Genbank accession number NC\_001802), such as to inhibit CD4 receptor mediated fusion of HIV-1. In particular, the present invention describes the selection and function of siNA molecules capable of modulating HIV-1 envelope glycoprotein expression, for example, expression of the gp120 and gp41 subunits of HIV-1 envelope glycoprotein. These siNA molecules can be used to treat diseases and disorders associated with HIV infection, or as a prophylactic measure to prevent HIV-1 infection.

In one embodiment, the invention features one or more siNA molecules and methods that independently or in combination modulate the expression of genes representing cellular targets for HIV infection, such as cellular receptors, cell surface molecules, cellular enzymes, cellular transcription factors, and/or cytokines, second messengers, and cellular accessory molecules.

Examples of such cellular receptors involved in HIV infection contemplated by the instant invention include, but are not limited to, CD4 receptors, CXCR4 (also known as Fusin; LESTR; NPY3R, e.g., Genbank Accession No. NM\_003467); CCR5 (also known as CKR-5, CMKRB5, e.g., Genbank Accession No. NM\_000579); CCR3 (also known as CC-CKR-3, CKR-3, CMKBR3, e.g., Genbank Accession No. NM\_001837); CCR2 (also known as CCR2b, CMKBR2, e.g., Genbank Accession Nos. NM\_000647 and

NM\_000648); CCR1 (also known as CKR1, CMKBR1, e.g., Genbank Accession No. NM\_001295); CCR4 (also known as CKR-4, e.g., Genbank Accession No. NM\_005508); CCR8 (also known as ChemR1, TER1, CMKBR8, e.g., Genbank Accession No. NM\_005201); CCR9 (also known as D6, e.g. Genbank Accession Nos. 5 NM\_006641 and NM\_031200); CXCR2 (also known as IL-8RB, e.g., Genbank Accession No. NM\_001557); STRL33 (also known as Bonzo; TYMSTR, e.g., Genbank Accession No. NM\_006564); US28; V28 (also known as CMKBRL1, CX3CR1, GPR13, e.g., Genbank Accession No. NM\_001337); gpr1 (also known as GPR1, e.g., Genbank Accession No. NM\_005279); gpr15 (also known as BOB, GPR15, e.g., Genbank 10 Accession No. NM\_005290); Apj (also known as angiotensin-receptor-like, AGTRL1, e.g., Genbank Accession No. NM\_005161); and ChemR23 receptors (e.g., Genbank Accession No. NM\_004072).

Examples of cell surface molecules involved in HIV infection contemplated by the instant invention include, but are not limited to, Heparan Sulfate Proteoglycans, HSPG2 15 (e.g., Genbank Accession No. NM\_005529); SDC2 (e.g., Genbank Accession Nos. AK025488, J04621, J04621); SDC4 (e.g., Genbank Accession No. NM\_002999); GPC1 (e.g., Genbank Accession No. NM\_002081); SDC3 (e.g., Genbank Accession No. NM\_014654); SDC1 (e.g., Genbank Accession No. NM\_002997); Galactoceramides (e.g., Genbank Accession Nos. NM\_000153, NM\_003360, NM\_001478.2, NM\_004775, 20 and NM\_004861); and Erythrocyte-expressed Glycolipids (e.g., Genbank Accession Nos. NM\_003778, NM\_003779, NM\_003780, NM\_030587, and NM\_001497).

Examples of cellular enzymes involved in HIV infection contemplated by the invention include, but are not limited to, N-myristoyltransferase (NMT1, e.g., Genbank Accession No. NM\_021079 and NMT2, e.g., Genbank Accession No. NM\_004808); 25 Glycosylation Enzymes (e.g., Genbank Accession Nos. NM\_000303, NM\_013339, NM\_003358, NM\_005787, NM\_002408, NM\_002676, NM\_002435), NM\_002409, NM\_006122, NM\_002372, NM\_006699, NM\_005907, NM\_004479, NM\_000150, NM\_005216 and NM\_005668); gp-160 Processing Enzymes (such as PCSK5, e.g., Genbank Accession No. NM\_006200); Ribonucleotide Reductase (e.g., Genbank 30 Accession Nos. NM\_001034, NM\_001033, AB036063, AB036063, AB036532, AK001965, AK001965, AK023605, AL137348, and AL137348); and Polyamine

Biosynthesis enzymes (e.g., Genbank Accession Nos. NM\_002539, NM\_003132 and NM\_001634).

Examples of cellular transcription factors involved in HIV infection contemplated by the invention include, but are not limited to, SP-1 and NF-kappa B (such as NFKB2, e.g., Genbank Accession No. NM\_002502; RELA, e.g., Genbank Accession No. NM\_021975; and NFKB1, e.g., Genbank Accession No. NM\_003998).

Examples of cytokines and second messengers involved in HIV infection contemplated by the invention include, but are not limited to, Tumor Necrosis Factor- $\alpha$  (TNF- $\alpha$ , e.g., Genbank Accession No. NM\_000594); Interleukin 1 $\alpha$  (IL-1 $\alpha$ , e.g., Genbank Accession No. NM\_000575); Interleukin 6 (IL-6, e.g., Genbank Accession No. NM\_000600); Phospholipase C (PLC, e.g., Genbank Accession No. NM\_000933); and Protein Kinase C (PKC, e.g., Genbank Accession No. NM\_006255).

Examples of cellular accessory molecules involved in HIV infection contemplated by the invention include, but are not limited to, Cyclophilins, (such as PPID, e.g., Genbank Accession No. NM\_005038; PPIA, e.g., Genbank Accession No. NM\_021130; PPIE, e.g., Genbank Accession No. NM\_006112; PPIB, e.g., Genbank Accession No. NM\_000942; PPIF, e.g., Genbank Accession No. NM\_005729; PPIG, e.g., Genbank Accession No. NM\_004792; and PPIC, e.g., Genbank Accession No. NM\_000943); Mitogen Activated Protein Kinase (MAP-Kinase, such as MAPK1, e.g., Genbank Accession Nos. NM\_002745 and NM\_138957); and Extracellular Signal-Regulated Kinase (ERK-Kinase).

A siNA molecule can be adapted for use to treat HIV infection or acquired immunodeficiency syndrome (AIDS). A siNA molecule can comprise a sense region and an antisense region and wherein said antisense region comprises sequence complementary to a HIV RNA sequence and the sense region comprises sequence complementary to the antisense region. A siNA molecule can be assembled from two nucleic acid fragments wherein one fragment comprises the sense region and the second fragment comprises the antisense region of said siNA molecule. The sense region and antisense region can be connected via a linker molecule, including covalently connected via the linker molecule. The linker molecule can be a polynucleotide linker or a non-nucleotide linker.

In one embodiment, the invention features a siNA molecule having RNAi activity against HIV-1 RNA, wherein the siNA molecule comprises a sequence complementary to any RNA having HIV-1 encoding sequence, for example, Genbank Accession No. AJ302647. In another embodiment, the invention features a siNA molecule having RNAi activity against HIV-2 RNA, wherein the siNA molecule comprises a sequence complementary to any RNA having HIV-2 encoding sequence, for example Genbank Accession No. NC\_001722. In another embodiment, the invention features a siNA molecule having RNAi activity against FIV-1 RNA, wherein the siNA molecule comprises a sequence complementary to any RNA having FIV-1 encoding sequence, for example, Genbank Accession No. NC\_001482. In another embodiment, the invention features a siNA molecule having RNAi activity against SIV-1 RNA, wherein the siNA molecule comprises a sequence complementary to any RNA having SIV-1 encoding sequence, for example, Genbank Accession No. M66437.

In yet another embodiment, the invention features a siNA molecule comprising a sequence complementary to a sequence comprising Genbank Accession Nos. AJ302647 (HIV-1), NC\_001722 (HIV-2), NC\_001482 (FIV-1) and/or M66437 (SIV-1).

The description below of the various aspects and embodiments is provided with reference to the exemplary HIV-1 gene, referred to as HIV. However, such reference is meant to be exemplary only and the various aspects and embodiments described herein are also directed to other genes which encode HIV polypeptides and/or polypeptides of similar viruses to HIV, as well as cellular targets for HIV, such as those described herein. The various aspects and embodiments are also directed to other genes that are involved in the progression, development, or maintenance of human immunodeficiency virus (HIV) infection and acquired immunodeficiency syndrome (AIDS). Those additional genes can be analyzed for target sites using the methods described for HIV herein. Thus, the inhibition and the effects of such inhibition of the other genes can be performed as described herein. In other words, the term "HIV" as it is defined herein below and recited in the described embodiments, is meant to encompass genes associated with the development or maintenance of human immunodeficiency virus (HIV) infection and acquired immunodeficiency syndrome (AIDS), such as genes which encode HIV polypeptides and/or polypeptides of similar viruses to HIV genes, as well as cellular genes involved in HIV pathways of gene expression and/or HIV activity. Also, the term

“HIV” as it is defined herein below and recited in the described embodiments, is meant to encompass HIV viral gene products and cellular gene products involved in HIV infection, such as those described herein. Thus, each of the embodiments described herein with reference to the term “HIV” are applicable to all of the virus, cellular and  
5 viral protein, peptide, polypeptide, and/or polynucleotide molecules covered by the term “HIV”, as that term is defined herein.

Due to the high sequence variability of the HIV genome, selection of nucleic acid molecules for broad therapeutic applications would likely involve the conserved regions of the HIV genome. Specifically, the present invention describes nucleic acid molecules  
10 that cleave the conserved regions of the HIV genome. Therefore, one nucleic acid molecule can be designed to target all the different isolates of HIV. Nucleic acid molecules designed to target conserved regions of various HIV isolates can enable efficient inhibition of HIV replication in diverse subject populations and can ensure the effectiveness of the nucleic acid molecules against HIV quasi species which evolve due  
15 to mutations in the non-conserved regions of the HIV genome. Therefore a single siNA molecule can be targeted against all isolates of HIV by designing the siNA molecule to interact with conserved nucleotide sequences of HIV (such conserved sequences are expected to be present in the RNA of all HIV isolates).

In one embodiment, the invention features a siNA molecule that down-regulates  
20 expression of a HIV gene, for example, wherein the HIV gene comprises HIV encoding sequence.

In one embodiment, the invention features a siNA molecule having RNAi activity against HIV RNA, wherein the siNA molecule comprises a sequence complementary to any RNA having HIV or other encoding sequence, such as the exemplary sequences  
25 having GenBank Accession Nos. disclosed herein. Chemical modifications as shown in **Tables III and IV** or otherwise described herein can be applied to any siNA construct of the invention.

In another embodiment, the invention features a siNA molecule having RNAi activity against a HIV gene or related genes involved in HIV infection and AIDS,  
30 wherein the siNA molecule comprises nucleotide sequence complementary to the nucleotide sequence of such genes, such as the exemplary sequences having GenBank

Accession Nos. disclosed herein. In another embodiment, a siNA molecule of the invention includes nucleotide sequence that can interact with nucleotide sequence of a HIV gene and thereby mediate silencing of HIV gene expression, for example, wherein the siNA mediates regulation of HIV gene expression by cellular processes that modulate the chromatin structure of the HIV gene and prevent transcription of the HIV gene.

In another embodiment, the invention features a siNA molecule comprising nucleotide sequence, for example, nucleotide sequence in the antisense region of the siNA molecule that is complementary to a nucleotide sequence or portion of sequence of a HIV gene. In another embodiment, the invention features a siNA molecule comprising a region, for example, the antisense region of the siNA construct, complementary to a sequence or portion of sequence comprising a HIV gene sequence.

In one embodiment, the antisense region of HIV siNA constructs can comprise a sequence complementary to sequence having any of SEQ ID NOs. 1-738 or 1477-1482. In one embodiment, the antisense region can also comprise sequence having any of SEQ ID NOs. 739-1476, 1491-1498, 1507-1514, 1523-1530, 1535-1538, 1547-1554, 1557-1558, 1567-1574, 1595, 1597, 1599, 1601, 1603, or 1604. In another embodiment, the sense region of HIV constructs can comprise sequence having any of SEQ ID NOs. 1-738, 1477-1490, 1499-1506, 1515-1522, 1531-1534, 1539-1546, 1555-1556, 1559-1566, 1575-1576, 1594, 1596, 1598, 1600, or 1602. The sense region can comprise a sequence of SEQ ID NO. 1583 and the antisense region can comprise a sequence of SEQ ID NO. 1584. The sense region can comprise a sequence of SEQ ID NO. 1585 and the antisense region can comprise a sequence of SEQ ID NO. 1586. The sense region can comprise a sequence of SEQ ID NO. 1587 and the antisense region can comprise a sequence of SEQ ID NO. 1588. The sense region can comprise a sequence of SEQ ID NO. 1589 and the antisense region can comprise a sequence of SEQ ID NO. 1590. The sense region can comprise a sequence of SEQ ID NO. 1591 and the antisense region can comprise a sequence of SEQ ID NO. 1592. The sense region can comprise a sequence of SEQ ID NO. 1589 and the antisense region can comprise a sequence of SEQ ID NO. 1593.

In one embodiment, a siNA molecule of the invention comprises any of SEQ ID NOs. 1-1604. The sequences shown in SEQ ID NOs: 1-1604 are not limiting. A siNA

molecule of the invention can comprise any contiguous HIV sequence (e.g., about 19 to about 25, or about 19, 20, 21, 22, 23, 24 or 25 contiguous HIV nucleotides).

In yet another embodiment, the invention features a siNA molecule comprising a sequence, for example, the antisense sequence of the siNA construct, complementary to a  
5 sequence or portion of sequence comprising sequence represented by GenBank Accession Nos. disclosed herein. Chemical modifications in **Tables III and IV** and described herein can be applied to any siRNA construct of the invention.

In one embodiment of the invention a siNA molecule comprises an antisense strand having about 19 to about 29 nucleotides, wherein the antisense strand is  
10 complementary to a RNA sequence encoding a HIV protein, and wherein said siNA further comprises a sense strand having about 19 to about 29 (e.g., about 19, 20, 21, 22, 23, 24, 25, 26, 27, 28 or 29) nucleotides, and wherein said sense strand and said antisense strand are distinct nucleotide sequences with at least about 19 complementary nucleotides.

15 In another embodiment of the invention a siNA molecule of the invention comprises an antisense region having about 19 to about 29 (e.g., about 19, 20, 21, 22, 23, 24, 25, 26, 27, 28 or 29) nucleotides, wherein the antisense region is complementary to a RNA sequence encoding a HIV protein, and wherein said siNA further comprises a sense region having about 19 to about 29 nucleotides, wherein said sense region and said  
20 antisense region comprise a linear molecule with at least about 19 complementary nucleotides.

In one embodiment of the invention a siNA molecule comprises an antisense strand comprising a nucleotide sequence that is complementary to a nucleotide sequence or a portion thereof encoding a HIV protein. The siNA further comprises a sense strand,  
25 wherein said sense strand comprises a nucleotide sequence of a HIV gene or a portion thereof.

In another embodiment, a siNA molecule comprises an antisense region comprising a nucleotide sequence that is complementary to a nucleotide sequence or a portion thereof encoding a HIV protein. The siNA molecule further comprises a sense

region, wherein said sense region comprises a nucleotide sequence of a HIV gene or a portion thereof.

In one embodiment, a siNA molecule of the invention has RNAi activity that modulates expression of RNA encoded by a HIV gene. Because HIV genes can share some degree of sequence homology with each other, siNA molecules can be designed to target a class of HIV genes (and associated receptor or ligand genes) or alternately specific HIV genes by selecting sequences that are either shared amongst different HIV targets or alternatively that are unique for a specific HIV target. Therefore, in one embodiment, the siNA molecule can be designed to target conserved regions of HIV RNA sequence having homology between several HIV genes so as to target several HIV genes (e.g., different HIV strains, splice variants, mutant genes. etc.) with one siNA molecule. In another embodiment, the siNA molecule can be designed to target a sequence that is unique to a specific HIV RNA sequence due to the high degree of specificity that the siNA molecule requires to mediate RNAi activity.

In one embodiment, nucleic acid molecules of the invention that act as mediators of the RNA interference gene silencing response are double-stranded nucleic acid molecules. In another embodiment, the siNA molecules of the invention consist of duplexes containing about 19 base pairs between oligonucleotides comprising about 19 to about 25 (e.g., about 19, 20, 21, 22, 23, 24 or 25) nucleotides. In yet another embodiment, siNA molecules of the invention comprise duplexes with overhanging ends of about about 1 to about 3 (e.g., about 1, 2, or 3) nucleotides, for example about 21-nucleotide duplexes with about 19 base pairs and 3'-terminal mononucleotide, dinucleotide, or trinucleotide overhangs.

In one embodiment, the invention features one or more chemically-modified siNA constructs having specificity for HIV expressing nucleic acid molecules, such as RNA encoding a HIV protein. Non-limiting examples of such chemical modifications include without limitation phosphorothioate internucleotide linkages, 2'-deoxyribonucleotides, 2'-O-methyl ribonucleotides, 2'-deoxy-2'-fluoro ribonucleotides, "universal base" nucleotides, "acyclic" nucleotides, 5-C-methyl nucleotides, and terminal glyceryl and/or inverted deoxy abasic residue incorporation. These chemical modifications, when used in various siNA constructs, are shown to preserve RNAi activity in cells while at the



same time, dramatically increasing the serum stability of these compounds. Furthermore, contrary to the data published by Parrish *et al.*, *supra*, applicant demonstrates that multiple (greater than one) phosphorothioate substitutions are well-tolerated and confer substantial increases in serum stability for modified siNA  
5 constructs.

In one embodiment, a siNA molecule of the invention comprises modified nucleotides while maintaining the ability to mediate RNAi. The modified nucleotides can be used to improve *in vitro* or *in vivo* characteristics such as stability, activity, and/or bioavailability. For example, a siNA molecule of the invention can comprise modified  
10 nucleotides as a percentage of the total number of nucleotides present in the siNA molecule. As such, a siNA molecule of the invention can generally comprise about 5% to 100% modified nucleotides (e.g., 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% or 100% modified nucleotides). The actual percentage of modified nucleotides present in a given siNA molecule will  
15 depend on the total number of nucleotides present in the siNA. If the siNA molecule is single stranded, the percent modification can be based upon the total number of nucleotides present in the single stranded siNA molecules. Likewise, if the siNA molecule is double stranded, the percent modification can be based upon the total number of nucleotides present in the sense strand, antisense strand, or both the sense and  
20 antisense strands.

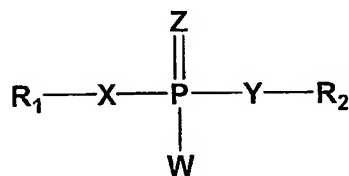
In a non-limiting example, the introduction of chemically-modified nucleotides into nucleic acid molecules provides a powerful tool in overcoming potential limitations of *in vivo* stability and bioavailability inherent to native RNA molecules that are delivered exogenously. For example, the use of chemically-modified nucleic acid  
25 molecules can enable a lower dose of a particular nucleic acid molecule for a given therapeutic effect since chemically-modified nucleic acid molecules tend to have a longer half-life in serum. Furthermore, certain chemical modifications can improve the bioavailability of nucleic acid molecules by targeting particular cells or tissues and/or improving cellular uptake of the nucleic acid molecule. Therefore, even if the activity of  
30 a chemically-modified nucleic acid molecule is reduced as compared to a native nucleic acid molecule, for example, when compared to an all-RNA nucleic acid molecule, the overall activity of the modified nucleic acid molecule can be greater than that of the

native molecule due to improved stability and/or delivery of the molecule. Unlike native unmodified siNA, chemically-modified siNA can also minimize the possibility of activating interferon activity in humans.

The antisense region of a siNA molecule of the invention can comprise a phosphorothioate internucleotide linkage at the 3'-end of said antisense region. The antisense region can comprise about one to about five phosphorothioate internucleotide linkages at the 5'-end of said antisense region. The 3'-terminal nucleotide overhangs of a siNA molecule of the invention can comprise ribonucleotides or deoxyribonucleotides that are chemically-modified at a nucleic acid sugar, base, or backbone. The 3'-terminal nucleotide overhangs can comprise one or more universal base ribonucleotides. The 3'-terminal nucleotide overhangs can comprise one or more acyclic nucleotides.

One embodiment of the invention provides an expression vector comprising a nucleic acid sequence encoding at least one siNA molecule of the invention in a manner that allows expression of the nucleic acid molecule. Another embodiment of the invention provides a mammalian cell comprising such an expression vector. The mammalian cell can be a human cell. The siNA molecule of the expression vector can comprise a sense region and an antisense region. The antisense region can comprise sequence complementary to a RNA or DNA sequence encoding HIV and the sense region can comprise sequence complementary to the antisense region. The siNA molecule can comprise two distinct strands having complementary sense and antisense regions. The siNA molecule can comprise a single strand having complementary sense and antisense regions.

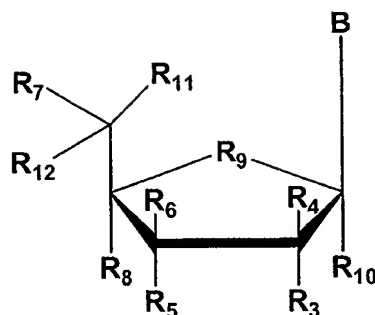
In one embodiment, the invention features a chemically-modified short interfering nucleic acid (siNA) molecule capable of mediating RNA interference (RNAi) against a HIV inside a cell or reconstituted *in vitro* system, wherein the chemical modification comprises one or more (*e.g.*, about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more) nucleotides comprising a backbone modified internucleotide linkage having Formula I:



wherein each R1 and R2 is independently any nucleotide, non-nucleotide, or polynucleotide which can be naturally-occurring or chemically-modified, each X and Y is independently O, S, N, alkyl, or substituted alkyl, each Z and W is independently O, S, N, alkyl, substituted alkyl, O-alkyl, S-alkyl, alkaryl, or aralkyl, and wherein W, X, Y, and Z are optionally not all O.

The chemically-modified internucleotide linkages having Formula I, for example, wherein any Z, W, X, and/or Y independently comprises a sulphur atom, can be present in one or both oligonucleotide strands of the siNA duplex, for example, in the sense strand, the antisense strand, or both strands. The siNA molecules of the invention can comprise one or more (*e.g.*, about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more) chemically-modified internucleotide linkages having Formula I at the 3'-end, the 5'-end, or both of the 3' and 5'-ends of the sense strand, the antisense strand, or both strands. For example, an exemplary siNA molecule of the invention can comprise about 1 to about 5 or more (*e.g.*, about 1, 2, 3, 4, 5, or more) chemically-modified internucleotide linkages having Formula I at the 5'-end of the sense strand, the antisense strand, or both strands. In another non-limiting example, an exemplary siNA molecule of the invention can comprise one or more (*e.g.*, about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more) pyrimidine nucleotides with chemically-modified internucleotide linkages having Formula I in the sense strand, the antisense strand, or both strands. In yet another non-limiting example, an exemplary siNA molecule of the invention can comprise one or more (*e.g.*, about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more) purine nucleotides with chemically-modified internucleotide linkages having Formula I in the sense strand, the antisense strand, or both strands. In another embodiment, a siNA molecule of the invention having internucleotide linkage(s) of Formula I also comprises a chemically-modified nucleotide or non-nucleotide having any of Formulae I-VII.

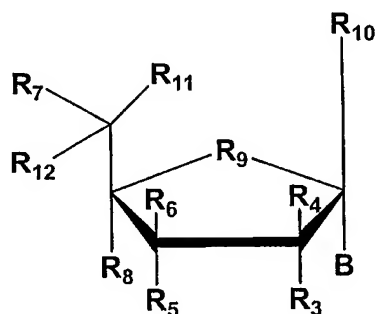
In one embodiment, the invention features a chemically-modified short interfering nucleic acid (siNA) molecule capable of mediating RNA interference (RNAi) against a HIV inside a cell or reconstituted *in vitro* system, wherein the chemical modification comprises one or more (*e.g.*, about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more) nucleotides or non-nucleotides having Formula II:



wherein each R3, R4, R5, R6, R7, R8, R10, R11 and R12 is independently H, OH, alkyl, substituted alkyl, alkaryl or aralkyl, F, Cl, Br, CN, CF3, OCF3, OCN, O-alkyl, S-alkyl, N-alkyl, O-alkenyl, S-alkenyl, N-alkenyl, SO-alkyl, alkyl-OSH, alkyl-OH, O-alkyl-OH, O-alkyl-SH, S-alkyl-OH, S-alkyl-SH, alkyl-S-alkyl, alkyl-O-alkyl, ONO2, NO2, N3, NH2, aminoalkyl, aminoacid, aminoacyl, ONH2, O-aminoalkyl, O-aminoacid, O-aminoacyl, heterocycloalkyl, heterocycloalkaryl, aminoalkylamino, polyalkylamino, substituted silyl, or group having Formula I; R9 is O, S, CH2, S=O, CHF, or CF2, and B is a nucleosidic base such as adenine, guanine, uracil, cytosine, thymine, 2-aminoadenosine, 5-methylcytosine, 2,6-diaminopurine, or any other non-naturally occurring base that can be complementary or non-complementary to target RNA or a non-nucleosidic base such as phenyl, naphthyl, 3-nitropyrrole, 5-nitroindole, nebularine, pyridone, pyridinone, or any other non-naturally occurring universal base that can be complementary or non-complementary to target RNA.

The chemically-modified nucleotide or non-nucleotide of Formula II can be present in one or both oligonucleotide strands of the siNA duplex, for example in the sense strand, the antisense strand, or both strands. The siNA molecules of the invention can comprise one or more chemically-modified nucleotide or non-nucleotide of Formula II at the 3'-end, the 5'-end, or both of the 3' and 5'-ends of the sense strand, the antisense strand, or both strands. For example, an exemplary siNA molecule of the invention can comprise about 1 to about 5 or more (*e.g.*, about 1, 2, 3, 4, 5, or more) chemically-modified nucleotides or non-nucleotides of Formula II at the 5'-end of the sense strand, the antisense strand, or both strands. In another non-limiting example, an exemplary siNA molecule of the invention can comprise about 1 to about 5 or more (*e.g.*, about 1, 2, 3, 4, 5, or more) chemically-modified nucleotides or non-nucleotides of Formula II at the 3'-end of the sense strand, the antisense strand, or both strands.

In one embodiment, the invention features a chemically-modified short interfering nucleic acid (siNA) molecule capable of mediating RNA interference (RNAi) against a HIV inside a cell or reconstituted *in vitro* system, wherein the chemical modification comprises one or more (*e.g.*, about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more) nucleotides or  
 5 non-nucleotides having Formula III:



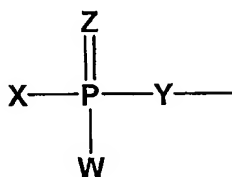
wherein each R3, R4, R5, R6, R7, R8, R10, R11 and R12 is independently H, OH, alkyl, substituted alkyl, alkaryl or aralkyl, F, Cl, Br, CN, CF3, OCF3, OCN, O-alkyl, S-alkyl, N-alkyl, O-alkenyl, S-alkenyl, N-alkenyl, SO-alkyl, alkyl-OSH, alkyl-OH, O-alkyl-OH,  
 10 O-alkyl-SH, S-alkyl-OH, S-alkyl-SH, alkyl-S-alkyl, alkyl-O-alkyl, ONO2, NO2, N3, NH2, aminoalkyl, aminoacid, aminoacyl, ONH2, O-aminoalkyl, O-aminoacid, O-aminoacyl, heterocycloalkyl, heterocycloalkaryl, aminoalkylamino, polyalkylamino, substituted silyl, or group having Formula I; R9 is O, S, CH2, S=O, CHF, or CF2, and B is a nucleosidic base such as adenine, guanine, uracil, cytosine, thymine, 2-aminoadenosine, 5-methylcytosine, 2,6-diaminopurine, or any other non-naturally  
 15 occurring base that can be employed to be complementary or non-complementary to target RNA or a non-nucleosidic base such as phenyl, naphthyl, 3-nitropyrrole, 5-nitroindole, nebularine, pyridone, pyridinone, or any other non-naturally occurring universal base that can be complementary or non-complementary to target RNA.

20 The chemically-modified nucleotide or non-nucleotide of Formula III can be present in one or both oligonucleotide strands of the siNA duplex, for example, in the sense strand, the antisense strand, or both strands. The siNA molecules of the invention can comprise one or more chemically-modified nucleotide or non-nucleotide of Formula III at the 3'-end, the 5'-end, or both of the 3' and 5'-ends of the sense strand, the antisense  
 25 strand, or both strands. For example, an exemplary siNA molecule of the invention can comprise about 1 to about 5 or more (*e.g.*, about 1, 2, 3, 4, 5, or more) chemically-

modified nucleotide(s) or non-nucleotide(s) of Formula III at the 5'-end of the sense strand, the antisense strand, or both strands. In another non-limiting example, an exemplary siNA molecule of the invention can comprise about 1 to about 5 or more (*e.g.*, about 1, 2, 3, 4, 5, or more) chemically-modified nucleotide or non-nucleotide of  
 5 Formula III, at the 3'-end of the sense strand, the antisense strand, or both strands.

In another embodiment, a siNA molecule of the invention comprises a nucleotide having Formula II or III, wherein the nucleotide having Formula II or III is in an inverted configuration. For example, the nucleotide having Formula II or III is connected to the siNA construct in a 3'-3', 3'-2', 2'-3', or 5'-5' configuration, such as at the 3'-end, the 5'-  
 10 end, or both of the 3' and 5'-ends of one or both siNA strands.

In one embodiment, the invention features a chemically-modified short interfering nucleic acid (siNA) molecule capable of mediating RNA interference (RNAi) against a HIV inside a cell or reconstituted *in vitro* system, wherein the chemical modification comprises a 5'-terminal phosphate group having Formula IV:



15

wherein each X and Y is independently O, S, N, alkyl, substituted alkyl, or alkylhalo; wherein each Z and W is independently O, S, N, alkyl, substituted alkyl, O-alkyl, S-alkyl, alkaryl, aralkyl, or alkylhalo; and wherein W, X, Y and Z are not all O.

In one embodiment, the invention features a siNA molecule having a 5'-terminal  
 20 phosphate group having Formula IV on the target-complementary strand, for example, a strand complementary to a target RNA, wherein the siNA molecule comprises an all RNA siNA molecule. In another embodiment, the invention features a siNA molecule having a 5'-terminal phosphate group having Formula IV on the target-complementary strand wherein the siNA molecule also comprises about 1 to about 3 (*e.g.*, about 1, 2, or  
 25 3) nucleotide 3'-terminal nucleotide overhangs having about 1 to about 4 (*e.g.*, about 1, 2, 3, or 4) deoxyribonucleotides on the 3'-end of one or both strands. In another embodiment, a 5'-terminal phosphate group having Formula IV is present on the target-

complementary strand of a siNA molecule of the invention, for example a siNA molecule having chemical modifications having any of Formulae I-VII.

In one embodiment, the invention features a chemically-modified short interfering nucleic acid (siNA) molecule capable of mediating RNA interference (RNAi) against a  
5 HIV inside a cell or reconstituted *in vitro* system, wherein the chemical modification comprises one or more phosphorothioate internucleotide linkages. For example, in a non-limiting example, the invention features a chemically-modified short interfering nucleic acid (siNA) having about 1, 2, 3, 4, 5, 6, 7, 8 or more phosphorothioate internucleotide linkages in one siNA strand. In yet another embodiment, the invention  
10 features a chemically-modified short interfering nucleic acid (siNA) individually having about 1, 2, 3, 4, 5, 6, 7, 8 or more phosphorothioate internucleotide linkages in both siNA strands. The phosphorothioate internucleotide linkages can be present in one or both oligonucleotide strands of the siNA duplex, for example in the sense strand, the antisense strand, or both strands. The siNA molecules of the invention can comprise one or more  
15 phosphorothioate internucleotide linkages at the 3'-end, the 5'-end, or both of the 3'- and 5'-ends of the sense strand, the antisense strand, or both strands. For example, an exemplary siNA molecule of the invention can comprise about 1 to about 5 or more (*e.g.*, about 1, 2, 3, 4, 5, or more) consecutive phosphorothioate internucleotide linkages at the 5'-end of the sense strand, the antisense strand, or both strands. In another non-limiting  
20 example, an exemplary siNA molecule of the invention can comprise one or more (*e.g.*, about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more) pyrimidine phosphorothioate internucleotide linkages in the sense strand, the antisense strand, or both strands. In yet another non-limiting example, an exemplary siNA molecule of the invention can comprise one or more (*e.g.*, about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more) purine phosphorothioate  
25 internucleotide linkages in the sense strand, the antisense strand, or both strands.

In one embodiment, the invention features a siNA molecule, wherein the sense strand comprises one or more, for example, about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more phosphorothioate internucleotide linkages, and/or one or more (*e.g.*, about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more) 2'-deoxy, 2'-O-methyl, 2'-deoxy-2'-fluoro, and/or one or more (*e.g.*,  
30 about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more) universal base modified nucleotides, and optionally a terminal cap molecule at the 3'-end, the 5'-end, or both of the 3'- and 5'-ends of the sense strand; and wherein the antisense strand comprises about 1 to about 10 or

more, specifically about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more phosphorothioate internucleotide linkages, and/or one or more (*e.g.*, about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more) 2'-deoxy, 2'-O-methyl, 2'-deoxy-2'-fluoro, and/or one or more (*e.g.*, about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more) universal base modified nucleotides, and optionally a  
5 terminal cap molecule at the 3'-end, the 5'-end, or both of the 3'- and 5'-ends of the antisense strand. In another embodiment, one or more, for example about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more, pyrimidine nucleotides of the sense and/or antisense siNA strand are chemically-modified with 2'-deoxy, 2'-O-methyl and/or 2'-deoxy-2'-fluoro nucleotides, with or without one or more, for example about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more,  
10 phosphorothioate internucleotide linkages and/or a terminal cap molecule at the 3'-end, the 5'-end, or both of the 3'- and 5'-ends, being present in the same or different strand.

In another embodiment, the invention features a siNA molecule, wherein the sense strand comprises about 1 to about 5, specifically about 1, 2, 3, 4, or 5 phosphorothioate internucleotide linkages, and/or one or more (*e.g.*, about 1, 2, 3, 4, 5, or more) 2'-deoxy,  
15 2'-O-methyl, 2'-deoxy-2'-fluoro, and/or one or more (*e.g.*, about 1, 2, 3, 4, 5, or more) universal base modified nucleotides, and optionally a terminal cap molecule at the 3'-end, the 5'-end, or both of the 3'- and 5'-ends of the sense strand; and wherein the antisense strand comprises about 1 to about 5 or more, specifically about 1, 2, 3, 4, 5, or more phosphorothioate internucleotide linkages, and/or one or more (*e.g.*, about 1, 2, 3, 4, 5, 6,  
20 7, 8, 9, 10 or more) 2'-deoxy, 2'-O-methyl, 2'-deoxy-2'-fluoro, and/or one or more (*e.g.*, about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more) universal base modified nucleotides, and optionally a terminal cap molecule at the 3'-end, the 5'-end, or both of the 3'- and 5'-ends of the antisense strand. In another embodiment, one or more, for example, about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more, pyrimidine nucleotides of the sense and/or antisense siNA  
25 strand are chemically-modified with 2'-deoxy, 2'-O-methyl and/or 2'-deoxy-2'-fluoro nucleotides, with or without about 1 to about 5 or more, for example about 1, 2, 3, 4, 5, or more phosphorothioate internucleotide linkages and/or a terminal cap molecule at the 3'-end, the 5'-end, or both of the 3'- and 5'-ends, being present in the same or different strand.

30 In one embodiment, the invention features a siNA molecule, wherein the antisense strand comprises one or more, for example, about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more phosphorothioate internucleotide linkages, and/or about one or more (*e.g.*, about 1, 2, 3,



4, 5, 6, 7, 8, 9, 10 or more) 2'-deoxy, 2'-O-methyl, 2'-deoxy-2'-fluoro, and/or one or more (e.g., about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more) universal base modified nucleotides, and optionally a terminal cap molecule at the 3'-end, the 5'-end, or both of the 3'- and 5'-ends of the sense strand; and wherein the antisense strand comprises about 1 to about 10 or more, specifically about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more phosphorothioate internucleotide linkages, and/or one or more (e.g., about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more) 2'-deoxy, 2'-O-methyl, 2'-deoxy-2'-fluoro, and/or one or more (e.g., about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more) universal base modified nucleotides, and optionally a terminal cap molecule at the 3'-end, the 5'-end, or both of the 3'- and 5'-ends of the antisense strand. In another embodiment, one or more, for example about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more pyrimidine nucleotides of the sense and/or antisense siNA strand are chemically-modified with 2'-deoxy, 2'-O-methyl and/or 2'-deoxy-2'-fluoro nucleotides, with or without one or more, for example about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more phosphorothioate internucleotide linkages and/or a terminal cap molecule at the 3'-end, the 5'-end, or both of the 3' and 5'-ends, being present in the same or different strand.

In another embodiment, the invention features a siNA molecule, wherein the antisense strand comprises about 1 to about 5 or more, specifically about 1, 2, 3, 4, 5 or more phosphorothioate internucleotide linkages, and/or one or more (e.g., about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more) 2'-deoxy, 2'-O-methyl, 2'-deoxy-2'-fluoro, and/or one or more (e.g., about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more) universal base modified nucleotides, and optionally a terminal cap molecule at the 3'-end, the 5'-end, or both of the 3'- and 5'-ends of the sense strand; and wherein the antisense strand comprises about 1 to about 5 or more, specifically about 1, 2, 3, 4, 5 or more phosphorothioate internucleotide linkages, and/or one or more (e.g., about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more) 2'-deoxy, 2'-O-methyl, 2'-deoxy-2'-fluoro, and/or one or more (e.g., about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more) universal base modified nucleotides, and optionally a terminal cap molecule at the 3'-end, the 5'-end, or both of the 3'- and 5'-ends of the antisense strand. In another embodiment, one or more, for example about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more pyrimidine nucleotides of the sense and/or antisense siNA strand are chemically-modified with 2'-deoxy, 2'-O-methyl and/or 2'-deoxy-2'-fluoro nucleotides, with or without about 1 to about 5, for example about 1, 2, 3, 4, 5 or more phosphorothioate internucleotide linkages and/or a terminal cap molecule at the 3'-end, the 5'-end, or both of the 3'- and 5'-ends, being present in the same or different strand.

In one embodiment, the invention features a chemically-modified short interfering nucleic acid (siNA) molecule having about 1 to about 5, specifically about 1, 2, 3, 4, 5 or more phosphorothioate internucleotide linkages in each strand of the siNA molecule.

In another embodiment, the invention features a siNA molecule comprising 2'-5' internucleotide linkages. The 2'-5' internucleotide linkage(s) can be at the 3'-end, the 5'-end, or both of the 3'- and 5'-ends of one or both siNA sequence strands. In addition, the 2'-5' internucleotide linkage(s) can be present at various other positions within one or both siNA sequence strands, for example, about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more including every internucleotide linkage of a pyrimidine nucleotide in one or both strands of the siNA molecule can comprise a 2'-5' internucleotide linkage, or about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more including every internucleotide linkage of a purine nucleotide in one or both strands of the siNA molecule can comprise a 2'-5' internucleotide linkage.

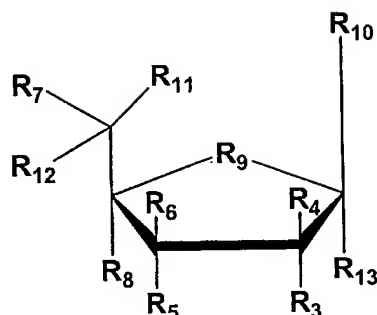
In another embodiment, a chemically-modified siNA molecule of the invention comprises a duplex having two strands, one or both of which can be chemically-modified, wherein each strand is about 18 to about 27 (*e.g.*, about 18, 19, 20, 21, 22, 23, 24, 25, 26, or 27) nucleotides in length, wherein the duplex has about 18 to about 23 (*e.g.*, about 18, 19, 20, 21, 22, or 23) base pairs, and wherein the chemical modification comprises a structure having any of Formulae I-VII. For example, an exemplary chemically-modified siNA molecule of the invention comprises a duplex having two strands, one or both of which can be chemically-modified with a chemical modification having any of Formulae I-VII or any combination thereof, wherein each strand consists of about 21 nucleotides, each having a 2-nucleotide 3'-terminal nucleotide overhang, and wherein the duplex has about 19 base pairs. In another embodiment, a siNA molecule of the invention comprises a single stranded hairpin structure, wherein the siNA is about 36 to about 70 (*e.g.*, about 36, 40, 45, 50, 55, 60, 65, or 70) nucleotides in length having about 18 to about 23 (*e.g.*, about 18, 19, 20, 21, 22, or 23) base pairs, and wherein the siNA can include a chemical modification comprising a structure having any of Formulae I-VII or any combination thereof. For example, an exemplary chemically-modified siNA molecule of the invention comprises a linear oligonucleotide having about 42 to about 50 (*e.g.*, about 42, 43, 44, 45, 46, 47, 48, 49, or 50) nucleotides that is chemically-modified with a chemical modification having any of Formulae I-VII or any combination thereof, wherein the linear oligonucleotide forms a hairpin structure having

about 19 base pairs and a 2-nucleotide 3'-terminal nucleotide overhang. In another embodiment, a linear hairpin siNA molecule of the invention contains a stem loop motif, wherein the loop portion of the siNA molecule is biodegradable. For example, a linear hairpin siNA molecule of the invention is designed such that degradation of the loop  
5 portion of the siNA molecule *in vivo* can generate a double-stranded siNA molecule with 3'-terminal overhangs, such as 3'-terminal nucleotide overhangs comprising about 2 nucleotides.

In another embodiment, a siNA molecule of the invention comprises a circular nucleic acid molecule, wherein the siNA is about 38 to about 70 (*e.g.*, about 38, 40, 45,  
10 50, 55, 60, 65, or 70) nucleotides in length having about 18 to about 23 (*e.g.*, about 18, 19, 20, 21, 22, or 23) base pairs, and wherein the siNA can include a chemical modification, which comprises a structure having any of Formulae I-VII or any combination thereof. For example, an exemplary chemically-modified siNA molecule of the invention comprises a circular oligonucleotide having about 42 to about 50 (*e.g.*,  
15 about 42, 43, 44, 45, 46, 47, 48, 49, or 50) nucleotides that is chemically-modified with a chemical modification having any of Formulae I-VII or any combination thereof, wherein the circular oligonucleotide forms a dumbbell shaped structure having about 19 base pairs and 2 loops.

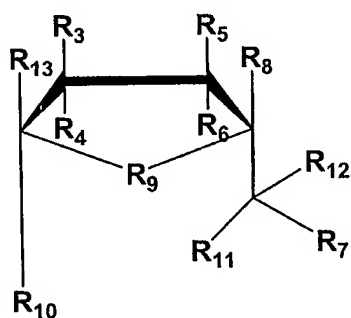
In another embodiment, a circular siNA molecule of the invention contains two  
20 loop motifs, wherein one or both loop portions of the siNA molecule is biodegradable. For example, a circular siNA molecule of the invention is designed such that degradation of the loop portions of the siNA molecule *in vivo* can generate a double-stranded siNA molecule with 3'-terminal overhangs, such as 3'-terminal nucleotide overhangs comprising about 2 nucleotides.

25 In one embodiment, a siNA molecule of the invention comprises at least one (*e.g.*, about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more) abasic moiety, for example a compound having Formula V:



wherein each R3, R4, R5, R6, R7, R8, R10, R11, R12, and R13 is independently H, OH, alkyl, substituted alkyl, alkaryl or aralkyl, F, Cl, Br, CN, CF3, OCF3, OCN, O-alkyl, S-alkyl, N-alkyl, O-alkenyl, S-alkenyl, N-alkenyl, SO-alkyl, alkyl-OSH, alkyl-OH, O-alkyl-OH, O-alkyl-SH, S-alkyl-OH, S-alkyl-SH, alkyl-S-alkyl, alkyl-O-alkyl, ONO2, NO2, N3, NH2, aminoalkyl, aminoacid, aminoacyl, ONH2, O-aminoalkyl, O-aminoacid, O-aminoacyl, heterocycloalkyl, heterocycloalkaryl, aminoalkylamino, polyalkylamino, substituted silyl, or group having Formula I; R9 is O, S, CH2, S=O, CHF, or CF2.

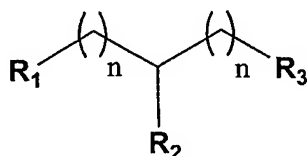
In one embodiment, a siNA molecule of the invention comprises at least one (*e.g.*, about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more) inverted abasic moiety, for example a compound having Formula VI:



wherein each R3, R4, R5, R6, R7, R8, R10, R11, R12, and R13 is independently H, OH, alkyl, substituted alkyl, alkaryl or aralkyl, F, Cl, Br, CN, CF3, OCF3, OCN, O-alkyl, S-alkyl, N-alkyl, O-alkenyl, S-alkenyl, N-alkenyl, SO-alkyl, alkyl-OSH, alkyl-OH, O-alkyl-OH, O-alkyl-SH, S-alkyl-OH, S-alkyl-SH, alkyl-S-alkyl, alkyl-O-alkyl, ONO2, NO2, N3, NH2, aminoalkyl, aminoacid, aminoacyl, ONH2, O-aminoalkyl, O-aminoacid, O-aminoacyl, heterocycloalkyl, heterocycloalkaryl, aminoalkylamino, polyalkylamino, substituted silyl, or group having Formula I; R9 is O, S, CH2, S=O, CHF, or CF2, and

either R2, R3, R8 or R13 serve as points of attachment to the siNA molecule of the invention.

In another embodiment, a siNA molecule of the invention comprises at least one (e.g., about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more) substituted polyalkyl moieties, for example a compound having Formula VII:



wherein each n is independently an integer from 1 to 12, each R1, R2 and R3 is independently H, OH, alkyl, substituted alkyl, alkaryl or aralkyl, F, Cl, Br, CN, CF3, OCF3, OCN, O-alkyl, S-alkyl, N-alkyl, O-alkenyl, S-alkenyl, N-alkenyl, SO-alkyl, alkyl-OSH, alkyl-OH, O-alkyl-OH, O-alkyl-SH, S-alkyl-OH, S-alkyl-SH, alkyl-S-alkyl, alkyl-O-alkyl, ONO2, NO2, N3, NH2, aminoalkyl, aminoacid, aminoacyl, ONH2, O-aminoalkyl, O-aminoacid, O-aminoacyl, heterocycloalkyl, heterocycloalkaryl, aminoalkylamino, polyalkylamino, substituted silyl, or a group having Formula I, and R1, R2 or R3 serves as points of attachment to the siNA molecule of the invention.

15 In another embodiment, the invention features a compound having Formula VII, wherein R1 and R2 are hydroxyl (OH) groups, n = 1, and R3 comprises O and is the point of attachment to the 3'-end, the 5'-end, or both of the 3' and 5'-ends of one or both strands of a double-stranded siNA molecule of the invention or to a single-stranded siNA molecule of the invention. This modification is referred to herein as "glyceryl" (for  
20 example modification 6 in **Figure 10**).

In another embodiment, a moiety having any of Formula V, VI or VII of the invention is at the 3'-end, the 5'-end, or both of the 3' and 5'-ends of a siNA molecule of the invention. For example, a moiety having Formula V, VI or VII can be present at the 3'-end, the 5'-end, or both of the 3' and 5'-ends of the antisense strand, the sense strand, or both antisense and sense strands of the siNA molecule. In addition, a moiety having Formula VII can be present at the 3'-end or the 5'-end of a hairpin siNA molecule as described herein.

In another embodiment, a siNA molecule of the invention comprises an abasic residue having Formula V or VI, wherein the abasic residue having Formula VI or VI is connected to the siNA construct in a 3'-3', 3'-2', 2'-3', or 5'-5' configuration, such as at the 3'-end, the 5'-end, or both of the 3' and 5'-ends of one or both siNA strands.

- 5 In one embodiment, a siNA molecule of the invention comprises one or more (*e.g.*, about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more) locked nucleic acid (LNA) nucleotides, for example at the 5'-end, the 3'-end, both of the 5' and 3'-ends, or any combination thereof, of the siNA molecule.

- 10 In another embodiment, a siNA molecule of the invention comprises one or more (*e.g.*, about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more) acyclic nucleotides, for example at the 5'-end, the 3'-end, both of the 5' and 3'-ends, or any combination thereof, of the siNA molecule.

- 15 In one embodiment, the invention features a chemically-modified short interfering nucleic acid (siNA) molecule of the invention, wherein the chemically-modified siNA comprises a sense region, where any (*e.g.*, one or more or all) pyrimidine nucleotides present in the sense region are 2'-deoxy-2'-fluoro pyrimidine nucleotides (*e.g.*, wherein all pyrimidine nucleotides are 2'-deoxy-2'-fluoro pyrimidine nucleotides or alternately a plurality of pyrimidine nucleotides are 2'-deoxy-2'-fluoro pyrimidine nucleotides), and where any (*e.g.*, one or more or all) purine nucleotides present in the sense region are 2'-  
20 deoxy purine nucleotides (*e.g.*, wherein all purine nucleotides are 2'-deoxy purine nucleotides or alternately a plurality of purine nucleotides are 2'-deoxy purine nucleotides).

- 25 In one embodiment, the invention features a chemically-modified short interfering nucleic acid (siNA) molecule of the invention, wherein the chemically-modified siNA comprises a sense region, where any (*e.g.*, one or more or all) pyrimidine nucleotides present in the sense region are 2'-deoxy-2'-fluoro pyrimidine nucleotides (*e.g.*, wherein all pyrimidine nucleotides are 2'-deoxy-2'-fluoro pyrimidine nucleotides or alternately a plurality of pyrimidine nucleotides are 2'-deoxy-2'-fluoro pyrimidine nucleotides), and where any (*e.g.*, one or more or all) purine nucleotides present in the sense region are 2'-  
30 deoxy purine nucleotides (*e.g.*, wherein all purine nucleotides are 2'-deoxy purine nucleotides or alternately a plurality of purine nucleotides are 2'-deoxy purine

nucleotides), wherein any nucleotides comprising a 3'-terminal nucleotide overhang that are present in said sense region are 2'-deoxy nucleotides.

In one embodiment, the invention features a chemically-modified short interfering nucleic acid (siNA) molecule of the invention, wherein the chemically-modified siNA  
5 comprises an antisense region, where any (*e.g.*, one or more or all) pyrimidine nucleotides present in the antisense region are 2'-deoxy-2'-fluoro pyrimidine nucleotides (*e.g.*, wherein all pyrimidine nucleotides are 2'-deoxy-2'-fluoro pyrimidine nucleotides or alternately a plurality of pyrimidine nucleotides are 2'-deoxy-2'-fluoro pyrimidine  
10 nucleotides), and wherein any (*e.g.*, one or more or all) purine nucleotides present in the antisense region are 2'-O-methyl purine nucleotides (*e.g.*, wherein all purine nucleotides are 2'-O-methyl purine nucleotides or alternately a plurality of purine nucleotides are 2'-O-methyl purine nucleotides).

In one embodiment, the invention features a chemically-modified short interfering nucleic acid (siNA) molecule of the invention, wherein the chemically-modified siNA  
15 comprises an antisense region, where any (*e.g.*, one or more or all) pyrimidine nucleotides present in the antisense region are 2'-deoxy-2'-fluoro pyrimidine nucleotides (*e.g.*, wherein all pyrimidine nucleotides are 2'-deoxy-2'-fluoro pyrimidine nucleotides or alternately a plurality of pyrimidine nucleotides are 2'-deoxy-2'-fluoro pyrimidine  
20 nucleotides), and wherein any (*e.g.*, one or more or all) purine nucleotides present in the antisense region are 2'-O-methyl purine nucleotides (*e.g.*, wherein all purine nucleotides are 2'-O-methyl purine nucleotides or alternately a plurality of purine nucleotides are 2'-O-methyl purine nucleotides), wherein any nucleotides comprising a 3'-terminal nucleotide overhang that are present in said antisense region are 2'-deoxy nucleotides.

In one embodiment, the invention features a chemically-modified short interfering  
25 nucleic acid (siNA) molecule of the invention, wherein the chemically-modified siNA comprises an antisense region, where any (*e.g.*, one or more or all) pyrimidine nucleotides present in the antisense region are 2'-deoxy-2'-fluoro pyrimidine nucleotides (*e.g.*, wherein all pyrimidine nucleotides are 2'-deoxy-2'-fluoro pyrimidine nucleotides or alternately a plurality of pyrimidine nucleotides are 2'-deoxy-2'-fluoro pyrimidine  
30 nucleotides), and where any (*e.g.*, one or more or all) purine nucleotides present in the antisense region are 2'-deoxy purine nucleotides (*e.g.*, wherein all purine nucleotides are

2'-deoxy purine nucleotides or alternately a plurality of purine nucleotides are 2'-deoxy purine nucleotides).

In one embodiment, the invention features a chemically-modified short interfering nucleic acid (siNA) molecule of the invention capable of mediating RNA interference (RNAi) against a HIV inside a cell or reconstituted *in vitro* system, wherein the chemically-modified siNA comprises a sense region, where one or more pyrimidine nucleotides present in the sense region are 2'-deoxy-2'-fluoro pyrimidine nucleotides (e.g., wherein all pyrimidine nucleotides are 2'-deoxy-2'-fluoro pyrimidine nucleotides or alternately a plurality of pyrimidine nucleotides are 2'-deoxy-2'-fluoro pyrimidine nucleotides), and where one or more purine nucleotides present in the sense region are 2'-deoxy purine nucleotides (e.g., wherein all purine nucleotides are 2'-deoxy purine nucleotides or alternately a plurality of purine nucleotides are 2'-deoxy purine nucleotides), and inverted deoxy abasic modifications that are optionally present at the 3'-end, the 5'-end, or both of the 3' and 5'-ends of the sense region, the sense region optionally further comprising a 3'-terminal overhang having about 1 to about 4 (e.g., about 1, 2, 3, or 4) 2'-deoxyribonucleotides; and wherein the chemically-modified short interfering nucleic acid molecule comprises an antisense region, where one or more pyrimidine nucleotides present in the antisense region are 2'-deoxy-2'-fluoro pyrimidine nucleotides (e.g., wherein all pyrimidine nucleotides are 2'-deoxy-2'-fluoro pyrimidine nucleotides or alternately a plurality of pyrimidine nucleotides are 2'-deoxy-2'-fluoro pyrimidine nucleotides), and wherein one or more purine nucleotides present in the antisense region are 2'-O-methyl purine nucleotides (e.g., wherein all purine nucleotides are 2'-O-methyl purine nucleotides or alternately a plurality of purine nucleotides are 2'-O-methyl purine nucleotides), and a terminal cap modification, such as any modification described herein or shown in **Figure 10**, that is optionally present at the 3'-end, the 5'-end, or both of the 3' and 5'-ends of the antisense sequence, the antisense region optionally further comprising a 3'-terminal nucleotide overhang having about 1 to about 4 (e.g., about 1, 2, 3, or 4) 2'-deoxynucleotides, wherein the overhang nucleotides can further comprise one or more (e.g., 1, 2, 3, or 4 ) phosphorothioate internucleotide linkages. Non-limiting examples of these chemically-modified siNAs are shown in **Figures 4 and 5** and **Tables III and IV** herein.



In one embodiment, the invention features a chemically-modified short interfering nucleic acid (siNA) molecule of the invention capable of mediating RNA interference (RNAi) against a HIV inside a cell or reconstituted *in vitro* system, wherein the siNA comprises a sense region, where one or more pyrimidine nucleotides present in the sense region are 2'-deoxy-2'-fluoro pyrimidine nucleotides (e.g., wherein all pyrimidine nucleotides are 2'-deoxy-2'-fluoro pyrimidine nucleotides or alternately a plurality of pyrimidine nucleotides are 2'-deoxy-2'-fluoro pyrimidine nucleotides), and where one or more purine nucleotides present in the sense region are purine ribonucleotides (e.g., wherein all purine nucleotides are purine ribonucleotides or alternately a plurality of purine nucleotides are purine ribonucleotides), and inverted deoxy abasic modifications that are optionally present at the 3'-end, the 5'-end, or both of the 3' and 5'-ends of the sense region, the sense region optionally further comprising a 3'-terminal overhang having about 1 to about 4 (e.g., about 1, 2, 3, or 4) 2'-deoxyribonucleotides; and wherein the siNA comprises an antisense region, where one or more pyrimidine nucleotides present in the antisense region are 2'-deoxy-2'-fluoro pyrimidine nucleotides (e.g., wherein all pyrimidine nucleotides are 2'-deoxy-2'-fluoro pyrimidine nucleotides or alternately a plurality of pyrimidine nucleotides are 2'-deoxy-2'-fluoro pyrimidine nucleotides), and wherein any purine nucleotides present in the antisense region are 2'-O-methyl purine nucleotides (e.g., wherein all purine nucleotides are 2'-O-methyl purine nucleotides or alternately a plurality of purine nucleotides are 2'-O-methyl purine nucleotides), and a terminal cap modification, such as any modification described herein or shown in **Figure 10**, that is optionally present at the 3'-end, the 5'-end, or both of the 3' and 5'-ends of the antisense sequence, the antisense region optionally further comprising a 3'-terminal nucleotide overhang having about 1 to about 4 (e.g., about 1, 2, 3, or 4) 2'-deoxynucleotides, wherein the overhang nucleotides can further comprise one or more (e.g., 1, 2, 3, or 4 ) phosphorothioate internucleotide linkages. Non-limiting examples of these chemically-modified siNAs are shown in **Figures 4 and 5** and **Tables III and IV** herein.

In one embodiment, the invention features a chemically-modified short interfering nucleic acid (siNA) molecule of the invention capable of mediating RNA interference (RNAi) against a HIV inside a cell or reconstituted *in vitro* system, wherein the chemically-modified siNA comprises a sense region, where one or more pyrimidine nucleotides present in the sense region are 2'-deoxy-2'-fluoro pyrimidine nucleotides

(e.g., wherein all pyrimidine nucleotides are 2'-deoxy-2'-fluoro pyrimidine nucleotides or alternately a plurality of pyrimidine nucleotides are 2'-deoxy-2'-fluoro pyrimidine nucleotides), and for example where one or more purine nucleotides present in the sense region are selected from the group consisting of 2'-deoxy nucleotides, locked nucleic acid (LNA) nucleotides, 2'-methoxyethyl nucleotides, 4'-thionucleotides, and 2'-O-methyl nucleotides (e.g., wherein all purine nucleotides are selected from the group consisting of 2'-deoxy nucleotides, locked nucleic acid (LNA) nucleotides, 2'-methoxyethyl nucleotides, 4'-thionucleotides, and 2'-O-methyl nucleotides or alternately a plurality of purine nucleotides are selected from the group consisting of 2'-deoxy nucleotides, locked nucleic acid (LNA) nucleotides, 2'-methoxyethyl nucleotides, 4'-thionucleotides, and 2'-O-methyl nucleotides), and wherein inverted deoxy abasic modifications are optionally present at the 3'-end, the 5'-end, or both of the 3' and 5'-ends of the sense region, the sense region optionally further comprising a 3'-terminal overhang having about 1 to about 4 (e.g., about 1, 2, 3, or 4) 2'-deoxyribonucleotides; and wherein the chemically-modified short interfering nucleic acid molecule comprises an antisense region, where one or more pyrimidine nucleotides present in the antisense region are 2'-deoxy-2'-fluoro pyrimidine nucleotides (e.g., wherein all pyrimidine nucleotides are 2'-deoxy-2'-fluoro pyrimidine nucleotides or alternately a plurality of pyrimidine nucleotides are 2'-deoxy-2'-fluoro pyrimidine nucleotides), and wherein one or more purine nucleotides present in the antisense region are selected from the group consisting of 2'-deoxy nucleotides, locked nucleic acid (LNA) nucleotides, 2'-methoxyethyl nucleotides, 4'-thionucleotides, and 2'-O-methyl nucleotides (e.g., wherein all purine nucleotides are selected from the group consisting of 2'-deoxy nucleotides, locked nucleic acid (LNA) nucleotides, 2'-methoxyethyl nucleotides, 4'-thionucleotides, and 2'-O-methyl nucleotides or alternately a plurality of purine nucleotides are selected from the group consisting of 2'-deoxy nucleotides, locked nucleic acid (LNA) nucleotides, 2'-methoxyethyl nucleotides, 4'-thionucleotides, and 2'-O-methyl nucleotides), and a terminal cap modification, such as any modification described herein or shown in **Figure 10**, that is optionally present at the 3'-end, the 5'-end, or both of the 3' and 5'-ends of the antisense sequence, the antisense region optionally further comprising a 3'-terminal nucleotide overhang having about 1 to about 4 (e.g., about 1, 2, 3, or 4) 2'-deoxynucleotides, wherein the overhang nucleotides can further comprise one or more (e.g., 1, 2, 3, or 4) phosphorothioate internucleotide linkages.

In another embodiment, any modified nucleotides present in the siNA molecules of the invention, preferably in the antisense strand of the siNA molecules of the invention, but also optionally in the sense and/or both antisense and sense strands, comprise modified nucleotides having properties or characteristics similar to naturally occurring  
5 ribonucleotides. For example, the invention features siNA molecules including modified nucleotides having a Northern conformation (e.g., Northern pseudorotation cycle, see for example Saenger, *Principles of Nucleic Acid Structure*, Springer-Verlag ed., 1984). As such, chemically modified nucleotides present in the siNA molecules of the invention, preferably in the antisense strand of the siNA molecules of the invention, but also  
10 optionally in the sense and/or both antisense and sense strands, are resistant to nuclease degradation while at the same time maintaining the capacity to mediate RNAi. Non-limiting examples of nucleotides having a northern configuration include locked nucleic acid (LNA) nucleotides (e.g., 2'-O,4'-C-methylene-(D-ribofuranosyl) nucleotides); 2'-methoxyethoxy (MOE) nucleotides; 2'-methyl-thio-ethyl, 2'-deoxy-2'-fluoro  
15 nucleotides, 2'-deoxy-2'-chloro nucleotides, 2'-azido nucleotides, and 2'-O-methyl nucleotides.

In one embodiment, the invention features a chemically-modified short interfering nucleic acid molecule (siNA) capable of mediating RNA interference (RNAi) against a HIV inside a cell or reconstituted *in vitro* system, wherein the chemical modification  
20 comprises a conjugate covalently attached to the chemically-modified siNA molecule. In another embodiment, the conjugate is covalently attached to the chemically-modified siNA molecule via a biodegradable linker. In one embodiment, the conjugate molecule is attached at the 3'-end of either the sense strand, the antisense strand, or both strands of the chemically-modified siNA molecule. In another embodiment, the conjugate  
25 molecule is attached at the 5'-end of either the sense strand, the antisense strand, or both strands of the chemically-modified siNA molecule. In yet another embodiment, the conjugate molecule is attached both the 3'-end and 5'-end of either the sense strand, the antisense strand, or both strands of the chemically-modified siNA molecule, or any combination thereof. In one embodiment, a conjugate molecule of the invention  
30 comprises a molecule that facilitates delivery of a chemically-modified siNA molecule into a biological system, such as a cell. In another embodiment, the conjugate molecule attached to the chemically-modified siNA molecule is a poly ethylene glycol, human serum albumin, or a ligand for a cellular receptor that can mediate cellular uptake.

Examples of specific conjugate molecules contemplated by the instant invention that can be attached to chemically-modified siNA molecules are described in Vargeese *et al.*, U.S. Serial No. 10/201,394, incorporated by reference herein. The type of conjugates used and the extent of conjugation of siNA molecules of the invention can be evaluated  
5 for improved pharmacokinetic profiles, bioavailability, and/or stability of siNA constructs while at the same time maintaining the ability of the siNA to mediate RNAi activity. As such, one skilled in the art can screen siNA constructs that are modified with various conjugates to determine whether the siNA conjugate complex possesses improved properties while maintaining the ability to mediate RNAi, for example in  
10 animal models as are generally known in the art.

In one embodiment, the invention features a short interfering nucleic acid (siNA) molecule of the invention, wherein the siNA further comprises a nucleotide, non-nucleotide, or mixed nucleotide/non-nucleotide linker that joins the sense region of the siNA to the antisense region of the siNA. In one embodiment, a nucleotide linker of the  
15 invention can be a linker of  $\geq 2$  nucleotides in length, for example 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides in length. In another embodiment, the nucleotide linker can be a nucleic acid aptamer. By "aptamer" or "nucleic acid aptamer" as used herein is meant a nucleic acid molecule that binds specifically to a target molecule wherein the nucleic acid molecule has sequence that comprises a sequence recognized by the target molecule in its natural  
20 setting. Alternately, an aptamer can be a nucleic acid molecule that binds to a target molecule where the target molecule does not naturally bind to a nucleic acid. The target molecule can be any molecule of interest. For example, the aptamer can be used to bind to a ligand-binding domain of a protein, thereby preventing interaction of the naturally occurring ligand with the protein. This is a non-limiting example and those in the art  
25 will recognize that other embodiments can be readily generated using techniques generally known in the art. (See, for example, Gold *et al.*, 1995, *Annu. Rev. Biochem.*, 64, 763; Brody and Gold, 2000, *J. Biotechnol.*, 74, 5; Sun, 2000, *Curr. Opin. Mol. Ther.*, 2, 100; Kusser, 2000, *J. Biotechnol.*, 74, 27; Hermann and Patel, 2000, *Science*, 287, 820; and Jayasena, 1999, *Clinical Chemistry*, 45, 1628.)

30 In yet another embodiment, a non-nucleotide linker of the invention comprises abasic nucleotide, polyether, polyamine, polyamide, peptide, carbohydrate, lipid, polyhydrocarbon, or other polymeric compounds (e.g. polyethylene glycols such as those

having between 2 and 100 ethylene glycol units). Specific examples include those described by Seela and Kaiser, *Nucleic Acids Res.* 1990, 18:6353 and *Nucleic Acids Res.* 1987, 15:3113; Cload and Schepartz, *J. Am. Chem. Soc.* 1991, 113:6324; Richardson and Schepartz, *J. Am. Chem. Soc.* 1991, 113:5109; Ma *et al.*, *Nucleic Acids Res.* 1993, 21:2585 and *Biochemistry* 1993, 32:1751; Durand *et al.*, *Nucleic Acids Res.* 1990, 18:6353; McCurdy *et al.*, *Nucleosides & Nucleotides* 1991, 10:287; Jschke *et al.*, *Tetrahedron Lett.* 1993, 34:301; Ono *et al.*, *Biochemistry* 1991, 30:9914; Arnold *et al.*, International Publication No. WO 89/02439; Usman *et al.*, International Publication No. WO 95/06731; Dudycz *et al.*, International Publication No. WO 95/11910 and Ferentz and Verdine, *J. Am. Chem. Soc.* 1991, 113:4000, all hereby incorporated by reference herein. A "non-nucleotide" further means any group or compound that can be incorporated into a nucleic acid chain in the place of one or more nucleotide units, including either sugar and/or phosphate substitutions, and allows the remaining bases to exhibit their enzymatic activity. The group or compound can be abasic in that it does not contain a commonly recognized nucleotide base, such as adenosine, guanine, cytosine, uracil or thymine, for example at the C1 position of the sugar.

In one embodiment, the invention features a short interfering nucleic acid (siNA) molecule capable of mediating RNA interference (RNAi) inside a cell or reconstituted in vitro system, wherein one or both strands of the siNA molecule that are assembled from two separate oligonucleotides do not comprise any ribonucleotides. For example, a siNA molecule can be assembled from a single oligonucleotide where the sense and antisense regions of the siNA comprise separate oligonucleotides not having any ribonucleotides (e.g., nucleotides having a 2'-OH group) present in the oligonucleotides. In another example, a siNA molecule can be assembled from a single oligonucleotide where the sense and antisense regions of the siNA are linked or circularized by a nucleotide or non-nucleotide linker as described herein, wherein the oligonucleotide does not have any ribonucleotides (e.g., nucleotides having a 2'-OH group) present in the oligonucleotide. Applicant has surprisingly found that the presense of ribonucleotides (e.g., nucleotides having a 2'-hydroxyl group) within the siNA molecule is not required or essential to support RNAi activity. As such, in one embodiment, all positions within the siNA can include chemically modified nucleotides and/or non-nucleotides such as nucleotides and or non-nucleotides having Formula I, II, III, IV, V, VI, or VII or any combination thereof

to the extent that the ability of the siNA molecule to support RNAi activity in a cell is maintained.

In one embodiment, a siNA molecule of the invention is a single stranded siNA molecule that mediates RNAi activity in a cell or reconstituted in vitro system, wherein the siNA molecule comprises a single stranded polynucleotide having complementarity to a target nucleic acid sequence. In another embodiment, the single stranded siNA molecule of the invention comprises a 5'-terminal phosphate group. In another embodiment, the single stranded siNA molecule of the invention comprises a 5'-terminal phosphate group and a 3'-terminal phosphate group (e.g., a 2',3'-cyclic phosphate). In another embodiment, the single stranded siNA molecule of the invention comprises about 19 to about 29 nucleotides. In yet another embodiment, the single stranded siNA molecule of the invention comprises one or more chemically modified nucleotides or non-nucleotides described herein. For example, all the positions within the siNA molecule can include chemically-modified nucleotides such as nucleotides having any of Formulae I-VII, or any combination thereof to the extent that the ability of the siNA molecule to support RNAi activity in a cell is maintained.

In one embodiment, a siNA molecule of the invention is a single stranded siNA molecule that mediates RNAi activity in a cell or reconstituted in vitro system, wherein the siNA molecule comprises a single stranded polynucleotide having complementarity to a target nucleic acid sequence, and wherein one or more pyrimidine nucleotides present in the siNA are 2'-deoxy-2'-fluoro pyrimidine nucleotides (e.g., wherein all pyrimidine nucleotides are 2'-deoxy-2'-fluoro pyrimidine nucleotides or alternately a plurality of pyrimidine nucleotides are 2'-deoxy-2'-fluoro pyrimidine nucleotides), and wherein any purine nucleotides present in the antisense region are 2'-O-methyl purine nucleotides (e.g., wherein all purine nucleotides are 2'-O-methyl purine nucleotides or alternately a plurality of purine nucleotides are 2'-O-methyl purine nucleotides), and a terminal cap modification, such as any modification described herein or shown in **Figure 10**, that is optionally present at the 3'-end, the 5'-end, or both of the 3' and 5'-ends of the antisense sequence, the siNA optionally further comprising about 1 to about 4 (e.g., about 1, 2, 3, or 4) terminal 2'-deoxynucleotides at the 3'-end of the siNA molecule, wherein the terminal nucleotides can further comprise one or more (e.g., 1, 2, 3, or 4 )

phosphorothioate internucleotide linkages, and wherein the siNA optionally further comprises a terminal phosphate group, such as a 5'-terminal phosphate group.

In one embodiment, a siNA molecule of the invention is a single stranded siNA molecule that mediates RNAi activity in a cell or reconstituted in vitro system, wherein the siNA molecule comprises a single stranded polynucleotide having complementarity  
5 to a target nucleic acid sequence, and wherein one or more pyrimidine nucleotides present in the siNA are 2'-deoxy-2'-fluoro pyrimidine nucleotides (e.g., wherein all pyrimidine nucleotides are 2'-deoxy-2'-fluoro pyrimidine nucleotides or alternately a plurality of pyrimidine nucleotides are 2'-deoxy-2'-fluoro pyrimidine nucleotides), and  
10 wherein any purine nucleotides present in the antisense region are 2'-deoxy purine nucleotides (e.g., wherein all purine nucleotides are 2'-deoxy purine nucleotides or alternately a plurality of purine nucleotides are 2'-deoxy purine nucleotides), and a terminal cap modification, such as any modification described herein or shown in **Figure 10**, that is optionally present at the 3'-end, the 5'-end, or both of the 3' and 5'-ends of the  
15 antisense sequence, the siNA optionally further comprising about 1 to about 4 (e.g., about 1, 2, 3, or 4) terminal 2'-deoxynucleotides at the 3'-end of the siNA molecule, wherein the terminal nucleotides can further comprise one or more (e.g., 1, 2, 3, or 4 ) phosphorothioate internucleotide linkages, and wherein the siNA optionally further comprises a terminal phosphate group, such as a 5'-terminal phosphate group.

20 In one embodiment, a siNA molecule of the invention is a single stranded siNA molecule that mediates RNAi activity in a cell or reconstituted in vitro system, wherein the siNA molecule comprises a single stranded polynucleotide having complementarity to a target nucleic acid sequence, and wherein one or more pyrimidine nucleotides present in the siNA are 2'-deoxy-2'-fluoro pyrimidine nucleotides (e.g., wherein all  
25 pyrimidine nucleotides are 2'-deoxy-2'-fluoro pyrimidine nucleotides or alternately a plurality of pyrimidine nucleotides are 2'-deoxy-2'-fluoro pyrimidine nucleotides), and wherein any purine nucleotides present in the antisense region are locked nucleic acid (LNA) nucleotides (e.g., wherein all purine nucleotides are LNA nucleotides or alternately a plurality of purine nucleotides are LNA nucleotides), and a terminal cap  
30 modification, such as any modification described herein or shown in **Figure 10**, that is optionally present at the 3'-end, the 5'-end, or both of the 3' and 5'-ends of the antisense sequence, the siNA optionally further comprising about 1 to about 4 (e.g., about 1, 2, 3,

or 4) terminal 2'-deoxynucleotides at the 3'-end of the siNA molecule, wherein the terminal nucleotides can further comprise one or more (e.g., 1, 2, 3, or 4 ) phosphorothioate internucleotide linkages, and wherein the siNA optionally further comprises a terminal phosphate group, such as a 5'-terminal phosphate group.

5        In one embodiment, a siNA molecule of the invention is a single stranded siNA molecule that mediates RNAi activity in a cell or reconstituted in vitro system, wherein the siNA molecule comprises a single stranded polynucleotide having complementarity to a target nucleic acid sequence, and wherein one or more pyrimidine nucleotides present in the siNA are 2'-deoxy-2'-fluoro pyrimidine nucleotides (e.g., wherein all  
10        pyrimidine nucleotides are 2'-deoxy-2'-fluoro pyrimidine nucleotides or alternately a plurality of pyrimidine nucleotides are 2'-deoxy-2'-fluoro pyrimidine nucleotides), and wherein any purine nucleotides present in the antisense region are 2'-methoxyethyl purine nucleotides (e.g., wherein all purine nucleotides are 2'-methoxyethyl purine nucleotides or alternately a plurality of purine nucleotides are 2'-methoxyethyl purine  
15        nucleotides), and a terminal cap modification, such as any modification described herein or shown in **Figure 10**, that is optionally present at the 3'-end, the 5'-end, or both of the 3' and 5'-ends of the antisense sequence, the siNA optionally further comprising about 1 to about 4 (e.g., about 1, 2, 3, or 4) terminal 2'-deoxynucleotides at the 3'-end of the siNA molecule, wherein the terminal nucleotides can further comprise one or more (e.g.,  
20        1, 2, 3, or 4 ) phosphorothioate internucleotide linkages, and wherein the siNA optionally further comprises a terminal phosphate group, such as a 5'-terminal phosphate group.

      In another embodiment, any modified nucleotides present in the single stranded siNA molecules of the invention comprise modified nucleotides having properties or characteristics similar to naturally occurring ribonucleotides. For example, the invention  
25        features siNA molecules including modified nucleotides having a Northern conformation (e.g., Northern pseudorotation cycle, see for example Saenger, *Principles of Nucleic Acid Structure*, Springer-Verlag ed., 1984). As such, chemically modified nucleotides present in the single stranded siNA molecules of the invention are preferably resistant to nuclease degradation while at the same time maintaining the capacity to mediate RNAi.

30        In one embodiment, the invention features a method for modulating the expression of a HIV gene within a cell comprising: (a) synthesizing a siNA molecule of the



invention, which can be chemically-modified, wherein one of the siNA strands comprises a sequence complementary to RNA of the HIV gene; and (b) introducing the siNA molecule into a cell under conditions suitable to modulate the expression of the HIV gene in the cell.

5           In one embodiment, the invention features a method for modulating the expression of a HIV gene within a cell comprising: (a) synthesizing a siNA molecule of the invention, which can be chemically-modified, wherein one of the siNA strands comprises a sequence complementary to RNA of the HIV gene and wherein the sense strand sequence of the siNA comprises a sequence identical to the sequence of the target  
10 RNA; and (b) introducing the siNA molecule into a cell under conditions suitable to modulate the expression of the HIV gene in the cell.

          In another embodiment, the invention features a method for modulating the expression of more than one HIV gene within a cell comprising: (a) synthesizing siNA molecules of the invention, which can be chemically-modified, wherein one of the siNA  
15 strands comprises a sequence complementary to RNA of the HIV genes; and (b) introducing the siNA molecules into a cell under conditions suitable to modulate the expression of the HIV genes in the cell.

          In another embodiment, the invention features a method for modulating the expression of more than one HIV gene within a cell comprising: (a) synthesizing a siNA  
20 molecule of the invention, which can be chemically-modified, wherein one of the siNA strands comprises a sequence complementary to RNA of the HIV gene and wherein the sense strand sequence of the siNA comprises a sequence identical to the sequence of the target RNA; and (b) introducing the siNA molecules into a cell under conditions suitable to modulate the expression of the HIV genes in the cell.

25           In one embodiment, siNA molecules of the invention are used as reagents in ex vivo applications. For example, siNA reagents are introduced into tissue or cells that are transplanted into a subject for therapeutic effect. The cells and/or tissue can be derived from an organism or subject that later receives the explant, or can be derived from another organism or subject prior to transplantation. The siNA molecules can be used to  
30 modulate the expression of one or more genes in the cells or tissue, such that the cells or tissue obtain a desired phenotype or are able to perform a function when transplanted in

vivo. In one embodiment, certain target cells from a patient are extracted. These extracted cells are contacted with siNAs targeting a specific nucleotide sequence within the cells under conditions suitable for uptake of the siNAs by these cells (e.g. using delivery reagents such as cationic lipids, liposomes and the like or using techniques such as electroporation to facilitate the delivery of siNAs into cells). The cells are then reintroduced back into the same patient or other patients.

In one embodiment, the invention features a method of modulating the expression of a HIV gene in a tissue explant comprising: (a) synthesizing a siNA molecule of the invention, which can be chemically-modified, wherein one of the siNA strands comprises a sequence complementary to RNA of the HIV gene; and (b) introducing the siNA molecule into a cell of the tissue explant derived from a particular organism under conditions suitable to modulate the expression of the HIV gene in the tissue explant. In another embodiment, the method further comprises introducing the tissue explant back into the organism the tissue was derived from or into another organism under conditions suitable to modulate the expression of the HIV gene in that organism.

In one embodiment, the invention features a method of modulating the expression of a HIV gene in a tissue explant comprising: (a) synthesizing a siNA molecule of the invention, which can be chemically-modified, wherein one of the siNA strands comprises a sequence complementary to RNA of the HIV gene and wherein the sense strand sequence of the siNA comprises a sequence identical to the sequence of the target RNA; and (b) introducing the siNA molecule into a cell of the tissue explant derived from a particular organism under conditions suitable to modulate the expression of the HIV gene in the tissue explant. In another embodiment, the method further comprises introducing the tissue explant back into the organism the tissue was derived from or into another organism under conditions suitable to modulate the expression of the HIV gene in that organism.

In another embodiment, the invention features a method of modulating the expression of more than one HIV gene in a tissue explant comprising: (a) synthesizing siNA molecules of the invention, which can be chemically-modified, wherein one of the siNA strands comprises a sequence complementary to RNA of the HIV genes; and (b) introducing the siNA molecules into a cell of the tissue explant derived from a particular

organism under conditions suitable to modulate the expression of the HIV genes in the tissue explant. In another embodiment, the method further comprises introducing the tissue explant back into the organism the tissue was derived from or into another organism under conditions suitable to modulate the expression of the HIV genes in that  
5 organism.

In one embodiment, the invention features a method of modulating the expression of a HIV gene in an organism comprising: (a) synthesizing a siNA molecule of the invention, which can be chemically-modified, wherein one of the siNA strands comprises a sequence complementary to RNA of the HIV gene; and (b) introducing the  
10 siNA molecule into the organism under conditions suitable to modulate the expression of the HIV gene in the organism.

In another embodiment, the invention features a method of modulating the expression of more than one HIV gene in an organism comprising: (a) synthesizing siNA molecules of the invention, which can be chemically-modified, wherein one of the  
15 siNA strands comprises a sequence complementary to RNA of the HIV genes; and (b) introducing the siNA molecules into the organism under conditions suitable to modulate the expression of the HIV genes in the organism.

In one embodiment, the invention features a method for modulating the expression of a HIV gene within a cell comprising: (a) synthesizing a siNA molecule of the invention, which can be chemically-modified, wherein the siNA comprises a single  
20 stranded sequence having complementarity to RNA of the HIV gene; and (b) introducing the siNA molecule into a cell under conditions suitable to modulate the expression of the HIV gene in the cell.

In another embodiment, the invention features a method for modulating the expression of more than one HIV gene within a cell comprising: (a) synthesizing siNA  
25 molecules of the invention, which can be chemically-modified, wherein the siNA comprises a single stranded sequence having complementarity to RNA of the HIV gene; and (b) contacting the siNA molecule with a cell in vitro or in vivo under conditions suitable to modulate the expression of the HIV genes in the cell.

In one embodiment, the invention features a method of modulating the expression of a HIV gene in a tissue explant comprising: (a) synthesizing a siNA molecule of the invention, which can be chemically-modified, wherein the siNA comprises a single stranded sequence having complementarity to RNA of the HIV gene; and (b) contacting  
5 the siNA molecule with a cell of the tissue explant derived from a particular organism under conditions suitable to modulate the expression of the HIV gene in the tissue explant. In another embodiment, the method further comprises introducing the tissue explant back into the organism the tissue was derived from or into another organism under conditions suitable to modulate the expression of the HIV gene in that organism.

10 In another embodiment, the invention features a method of modulating the expression of more than one HIV gene in a tissue explant comprising: (a) synthesizing siNA molecules of the invention, which can be chemically-modified, wherein the siNA comprises a single stranded sequence having complementarity to RNA of the HIV gene; and (b) introducing the siNA molecules into a cell of the tissue explant derived from a  
15 particular organism under conditions suitable to modulate the expression of the HIV genes in the tissue explant. In another embodiment, the method further comprises introducing the tissue explant back into the organism the tissue was derived from or into another organism under conditions suitable to modulate the expression of the HIV genes in that organism.

20 In one embodiment, the invention features a method of modulating the expression of a HIV gene in an organism comprising: (a) synthesizing a siNA molecule of the invention, which can be chemically-modified, wherein the siNA comprises a single stranded sequence having complementarity to RNA of the HIV gene; and (b) introducing the siNA molecule into the organism under conditions suitable to modulate the  
25 expression of the HIV gene in the organism.

In another embodiment, the invention features a method of modulating the expression of more than one HIV gene in an organism comprising: (a) synthesizing siNA molecules of the invention, which can be chemically-modified, wherein the siNA comprises a single stranded sequence having complementarity to RNA of the HIV gene;  
30 and (b) introducing the siNA molecules into the organism under conditions suitable to modulate the expression of the HIV genes in the organism.

In one embodiment, the invention features a method of modulating the expression of a HIV gene in an organism comprising contacting the organism with a siNA molecule of the invention under conditions suitable to modulate the expression of the HIV gene in the organism.

- 5 In another embodiment, the invention features a method of modulating the expression of more than one HIV gene in an organism comprising contacting the organism with one or more siNA molecules of the invention under conditions suitable to modulate the expression of the HIV genes in the organism.

The siNA molecules of the invention can be designed to inhibit target (HIV) gene  
10 expression through RNAi targeting of a variety of RNA molecules. In one embodiment, the siNA molecules of the invention are used to target various RNAs corresponding to a target gene. Non-limiting examples of such RNAs include messenger RNA (mRNA), alternate RNA splice variants of target gene(s), post-transcriptionally modified RNA of target gene(s), pre-mRNA of target gene(s), and/or RNA templates. If alternate splicing  
15 produces a family of transcripts that are distinguished by usage of appropriate exons, the instant invention can be used to inhibit gene expression through the appropriate exons to specifically inhibit or to distinguish among the functions of gene family members. For example, a protein that contains an alternatively spliced transmembrane domain can be expressed in both membrane bound and secreted forms. Use of the invention to target  
20 the exon containing the transmembrane domain can be used to determine the functional consequences of pharmaceutical targeting of membrane bound as opposed to the secreted form of the protein. Non-limiting examples of applications of the invention relating to targeting these RNA molecules include therapeutic pharmaceutical applications, pharmaceutical discovery applications, molecular diagnostic and gene function  
25 applications, and gene mapping, for example using single nucleotide polymorphism mapping with siNA molecules of the invention. Such applications can be implemented using known gene sequences or from partial sequences available from an expressed sequence tag (EST).

In another embodiment, the siNA molecules of the invention are used to target  
30 conserved sequences corresponding to a gene family or gene families such as HIV family genes. As such, siNA molecules targeting multiple HIV targets can provide increased

therapeutic effect. In addition, siNA can be used to characterize pathways of gene function in a variety of applications. For example, the present invention can be used to inhibit the activity of target gene(s) in a pathway to determine the function of uncharacterized gene(s) in gene function analysis, mRNA function analysis, or translational analysis. The invention can be used to determine potential target gene pathways involved in various diseases and conditions toward pharmaceutical development. The invention can be used to understand pathways of gene expression involved in, for example, the progression and/or maintenance of cancer.

In one embodiment, siNA molecule(s) and/or methods of the invention are used to inhibit the expression of gene(s) that encode RNA referred to by Genbank Accession, for example HIV genes encoding RNA sequence(s) referred to herein by Genbank Accession number, for example Genbank Accession Nos. disclosed herein.

In one embodiment, the invention features a method comprising: (a) generating a library of siNA constructs having a predetermined complexity; and (b) assaying the siNA constructs of (a) above, under conditions suitable to determine RNAi target sites within the target RNA sequence. In another embodiment, the siNA molecules of (a) have strands of a fixed length, for example, about 23 nucleotides in length. In yet another embodiment, the siNA molecules of (a) are of differing length, for example having strands of about 19 to about 25 (*e.g.*, about 19, 20, 21, 22, 23, 24, or 25) nucleotides in length. In one embodiment, the assay can comprise a reconstituted *in vitro* siNA assay as described herein. In another embodiment, the assay can comprise a cell culture system in which target RNA is expressed. In another embodiment, fragments of target RNA are analyzed for detectable levels of cleavage, for example by gel electrophoresis, northern blot analysis, or RNase protection assays, to determine the most suitable target site(s) within the target RNA sequence. The target RNA sequence can be obtained as is known in the art, for example, by cloning and/or transcription for *in vitro* systems, and by cellular expression in *in vivo* systems.

In one embodiment, the invention features a method comprising: (a) generating a randomized library of siNA constructs having a predetermined complexity, such as of  $4^N$ , where N represents the number of base paired nucleotides in each of the siNA construct strands (*eg.* for a siNA construct having 21 nucleotide sense and antisense strands with

19 base pairs, the complexity would be  $4^{19}$ ); and (b) assaying the siNA constructs of (a) above, under conditions suitable to determine RNAi target sites within the target HIV RNA sequence. In another embodiment, the siNA molecules of (a) have strands of a fixed length, for example about 23 nucleotides in length. In yet another embodiment, the siNA molecules of (a) are of differing length, for example having strands of about 19 to about 25 (e.g., about 19, 20, 21, 22, 23, 24, or 25) nucleotides in length. In one embodiment, the assay can comprise a reconstituted *in vitro* siNA assay as described in Example 7 herein. In another embodiment, the assay can comprise a cell culture system in which target RNA is expressed. In another embodiment, fragments of HIV RNA are analyzed for detectable levels of cleavage, for example by gel electrophoresis, northern blot analysis, or RNase protection assays, to determine the most suitable target site(s) within the target HIV RNA sequence. The target HIV RNA sequence can be obtained as is known in the art, for example, by cloning and/or transcription for *in vitro* systems, and by cellular expression in *in vivo* systems.

15 In another embodiment, the invention features a method comprising: (a) analyzing the sequence of a RNA target encoded by a target gene; (b) synthesizing one or more sets of siNA molecules having sequence complementary to one or more regions of the RNA of (a); and (c) assaying the siNA molecules of (b) under conditions suitable to determine RNAi targets within the target RNA sequence. In one embodiment, the siNA molecules of (b) have strands of a fixed length, for example about 23 nucleotides in length. In another embodiment, the siNA molecules of (b) are of differing length, for example having strands of about 19 to about 25 (e.g., about 19, 20, 21, 22, 23, 24, or 25) nucleotides in length. In one embodiment, the assay can comprise a reconstituted *in vitro* siNA assay as described herein. In another embodiment, the assay can comprise a cell culture system in which target RNA is expressed. Fragments of target RNA are analyzed for detectable levels of cleavage, for example by gel electrophoresis, northern blot analysis, or RNase protection assays, to determine the most suitable target site(s) within the target RNA sequence. The target RNA sequence can be obtained as is known in the art, for example, by cloning and/or transcription for *in vitro* systems, and by expression in *in vivo* systems.

By "target site" is meant a sequence within a target RNA that is "targeted" for cleavage mediated by a siNA construct which contains sequences within its antisense region that are complementary to the target sequence.

By "detectable level of cleavage" is meant cleavage of target RNA (and formation of cleaved product RNAs) to an extent sufficient to discern cleavage products above the background of RNAs produced by random degradation of the target RNA. Production of cleavage products from 1-5% of the target RNA is sufficient to detect above the background for most methods of detection.

In one embodiment, the invention features a composition comprising a siNA molecule of the invention, which can be chemically-modified, in a pharmaceutically acceptable carrier or diluent. In another embodiment, the invention features a pharmaceutical composition comprising siNA molecules of the invention, which can be chemically-modified, targeting one or more genes in a pharmaceutically acceptable carrier or diluent. In another embodiment, the invention features a method for treating or preventing a disease or condition in a subject, comprising administering to the subject a composition of the invention under conditions suitable for the treatment or prevention of the disease or condition in the subject, alone or in conjunction with one or more other therapeutic compounds. In yet another embodiment, the invention features a method for reducing or preventing tissue rejection in a subject comprising administering to the subject a composition of the invention under conditions suitable for the reduction or prevention of tissue rejection in the subject.

In another embodiment, the invention features a method for validating a HIV gene target comprising: (a) synthesizing a siNA molecule of the invention, which can be chemically-modified, wherein one of the siNA strands includes a sequence complementary to RNA of a HIV target gene; (b) introducing the siNA molecule into a cell, tissue, or organism under conditions suitable for modulating expression of the HIV target gene in the cell, tissue, or organism; and (c) determining the function of the gene by assaying for any phenotypic change in the cell, tissue, or organism.

In another embodiment, the invention features a method for validating a HIV target comprising: (a) synthesizing a siNA molecule of the invention, which can be chemically-modified, wherein one of the siNA strands includes a sequence



complementary to RNA of a HIV target gene; (b) introducing the siNA molecule into a biological system under conditions suitable for modulating expression of the HIV target gene in the biological system; and (c) determining the function of the gene by assaying for any phenotypic change in the biological system.

5 By "biological system" is meant, material, in a purified or unpurified form, from biological sources, including but not limited to human, animal, plant, insect, bacterial, viral or other sources, wherein the system comprises the components required for RNAi activity. The term "biological system" includes, for example, a cell, tissue, or organism, or extract thereof. The term biological system also includes reconstituted RNAi systems  
10 that can be used in an *in vitro* setting.

By "phenotypic change" is meant any detectable change to a cell that occurs in response to contact or treatment with a nucleic acid molecule of the invention (e.g., siNA). Such detectable changes include, but are not limited to, changes in shape, size, proliferation, motility, protein expression or RNA expression or other physical or  
15 chemical changes as can be assayed by methods known in the art. The detectable change can also include expression of reporter genes/molecules such as Green Florescent Protein (GFP) or various tags that are used to identify an expressed protein or any other cellular component that can be assayed.

In one embodiment, the invention features a kit containing a siNA molecule of the  
20 invention, which can be chemically-modified, that can be used to modulate the expression of a HIV target gene in a cell, tissue, or organism. In another embodiment, the invention features a kit containing more than one siNA molecule of the invention, which can be chemically-modified, that can be used to modulate the expression of more than one HIV target gene in a cell, tissue, or organism.

25 In one embodiment, the invention features a cell containing one or more siNA molecules of the invention, which can be chemically-modified. In another embodiment, the cell containing a siNA molecule of the invention is a mammalian cell. In yet another embodiment, the cell containing a siNA molecule of the invention is a human cell.

In one embodiment, the synthesis of a siNA molecule of the invention, which can  
30 be chemically-modified, comprises: (a) synthesis of two complementary strands of the

siNA molecule; (b) annealing the two complementary strands together under conditions suitable to obtain a double-stranded siNA molecule. In another embodiment, synthesis of the two complementary strands of the siNA molecule is by solid phase oligonucleotide synthesis. In yet another embodiment, synthesis of the two complementary strands of the  
5 siNA molecule is by solid phase tandem oligonucleotide synthesis.

In one embodiment, the invention features a method for synthesizing a siNA duplex molecule comprising: (a) synthesizing a first oligonucleotide sequence strand of the siNA molecule, wherein the first oligonucleotide sequence strand comprises a  
10 cleavable linker molecule that can be used as a scaffold for the synthesis of the second oligonucleotide sequence strand of the siNA; (b) synthesizing the second oligonucleotide sequence strand of siNA on the scaffold of the first oligonucleotide sequence strand, wherein the second oligonucleotide sequence strand further comprises a chemical moiety than can be used to purify the siNA duplex; (c) cleaving the linker molecule of (a) under  
15 conditions suitable for the two siNA oligonucleotide strands to hybridize and form a stable duplex; and (d) purifying the siNA duplex utilizing the chemical moiety of the second oligonucleotide sequence strand. In one embodiment, cleavage of the linker molecule in (c) above takes place during deprotection of the oligonucleotide, for example under hydrolysis conditions using an alkylamine base such as methylamine. In one  
20 embodiment, the method of synthesis comprises solid phase synthesis on a solid support such as controlled pore glass (CPG) or polystyrene, wherein the first sequence of (a) is synthesized on a cleavable linker, such as a succinyl linker, using the solid support as a scaffold. The cleavable linker in (a) used as a scaffold for synthesizing the second strand can comprise similar reactivity as the solid support derivatized linker, such that cleavage  
25 of the solid support derivatized linker and the cleavable linker of (a) takes place concomitantly. In another embodiment, the chemical moiety of (b) that can be used to isolate the attached oligonucleotide sequence comprises a trityl group, for example a dimethoxytrityl group, which can be employed in a trityl-on synthesis strategy as described herein. In yet another embodiment, the chemical moiety, such as a dimethoxytrityl group, is removed during purification, for example, using acidic  
30 conditions.

In a further embodiment, the method for siNA synthesis is a solution phase synthesis or hybrid phase synthesis wherein both strands of the siNA duplex are

synthesized in tandem using a cleavable linker attached to the first sequence which acts a scaffold for synthesis of the second sequence. Cleavage of the linker under conditions suitable for hybridization of the separate siNA sequence strands results in formation of the double-stranded siNA molecule.

5           In another embodiment, the invention features a method for synthesizing a siNA duplex molecule comprising: (a) synthesizing one oligonucleotide sequence strand of the siNA molecule, wherein the sequence comprises a cleavable linker molecule that can be used as a scaffold for the synthesis of another oligonucleotide sequence; (b) synthesizing a second oligonucleotide sequence having complementarity to the first  
10   sequence strand on the scaffold of (a), wherein the second sequence comprises the other strand of the double-stranded siNA molecule and wherein the second sequence further comprises a chemical moiety than can be used to isolate the attached oligonucleotide sequence; (c) purifying the product of (b) utilizing the chemical moiety of the second oligonucleotide sequence strand under conditions suitable for isolating the full-length  
15   sequence comprising both siNA oligonucleotide strands connected by the cleavable linker and under conditions suitable for the two siNA oligonucleotide strands to hybridize and form a stable duplex. In one embodiment, cleavage of the linker molecule in (c) above takes place during deprotection of the oligonucleotide, for example under hydrolysis conditions. In another embodiment, cleavage of the linker molecule in (c)  
20   above takes place after deprotection of the oligonucleotide. In another embodiment, the method of synthesis comprises solid phase synthesis on a solid support such as controlled pore glass (CPG) or polystyrene, wherein the first sequence of (a) is synthesized on a cleavable linker, such as a succinyl linker, using the solid support as a scaffold. The cleavable linker in (a) used as a scaffold for synthesizing the second strand can comprise  
25   similar reactivity or differing reactivity as the solid support derivatized linker, such that cleavage of the solid support derivatized linker and the cleavable linker of (a) takes place either concomitantly or sequentially. In one embodiment, the chemical moiety of (b) that can be used to isolate the attached oligonucleotide sequence comprises a trityl group, for example a dimethoxytrityl group.

30           In another embodiment, the invention features a method for making a double-stranded siNA molecule in a single synthetic process comprising: (a) synthesizing an oligonucleotide having a first and a second sequence, wherein the first sequence is

complementary to the second sequence, and the first oligonucleotide sequence is linked to the second sequence via a cleavable linker, and wherein a terminal 5'-protecting group, for example, a 5'-O-dimethoxytrityl group (5'-O-DMT) remains on the oligonucleotide having the second sequence; (b) deprotecting the oligonucleotide whereby the  
5 deprotection results in the cleavage of the linker joining the two oligonucleotide sequences; and (c) purifying the product of (b) under conditions suitable for isolating the double-stranded siNA molecule, for example using a trityl-on synthesis strategy as described herein.

In another embodiment, the method of synthesis of siNA molecules of the  
10 invention comprises the teachings of Scaringe *et al.*, US Patent Nos. 5,889,136; 6,008,400; and 6,111,086, incorporated by reference herein in their entirety.

In one embodiment, the invention features siNA constructs that mediate RNAi against a HIV, wherein the siNA construct comprises one or more chemical modifications, for example, one or more chemical modifications having any of Formulae  
15 I-VII or any combination thereof that increases the nuclease resistance of the siNA construct.

In another embodiment, the invention features a method for generating siNA molecules with increased nuclease resistance comprising (a) introducing nucleotides having any of Formula I-VII or any combination thereof into a siNA molecule, and (b)  
20 assaying the siNA molecule of step (a) under conditions suitable for isolating siNA molecules having increased nuclease resistance.

In one embodiment, the invention features siNA constructs that mediate RNAi against a HIV, wherein the siNA construct comprises one or more chemical modifications described herein that modulates the binding affinity between the sense and  
25 antisense strands of the siNA construct.

In another embodiment, the invention features a method for generating siNA molecules with increased binding affinity between the sense and antisense strands of the siNA molecule comprising (a) introducing nucleotides having any of Formula I-VII or any combination thereof into a siNA molecule, and (b) assaying the siNA molecule of

step (a) under conditions suitable for isolating siNA molecules having increased binding affinity between the sense and antisense strands of the siNA molecule.

In one embodiment, the invention features siNA constructs that mediate RNAi against a HIV, wherein the siNA construct comprises one or more chemical  
5 modifications described herein that modulates the binding affinity between the antisense strand of the siNA construct and a complementary target RNA sequence within a cell.

In one embodiment, the invention features siNA constructs that mediate RNAi against a HIV, wherein the siNA construct comprises one or more chemical  
10 modifications described herein that modulates the binding affinity between the antisense strand of the siNA construct and a complementary target DNA sequence within a cell.

In another embodiment, the invention features a method for generating siNA molecules with increased binding affinity between the antisense strand of the siNA molecule and a complementary target RNA sequence comprising (a) introducing  
15 nucleotides having any of Formula I-VII or any combination thereof into a siNA molecule, and (b) assaying the siNA molecule of step (a) under conditions suitable for isolating siNA molecules having increased binding affinity between the antisense strand of the siNA molecule and a complementary target RNA sequence.

In another embodiment, the invention features a method for generating siNA molecules with increased binding affinity between the antisense strand of the siNA  
20 molecule and a complementary target DNA sequence comprising (a) introducing nucleotides having any of Formula I-VII or any combination thereof into a siNA molecule, and (b) assaying the siNA molecule of step (a) under conditions suitable for isolating siNA molecules having increased binding affinity between the antisense strand of the siNA molecule and a complementary target DNA sequence.

25 In one embodiment, the invention features siNA constructs that mediate RNAi against a HIV, wherein the siNA construct comprises one or more chemical modifications described herein that modulate the polymerase activity of a cellular polymerase capable of generating additional endogenous siNA molecules having sequence homology to the chemically-modified siNA construct.

In another embodiment, the invention features a method for generating siNA molecules capable of mediating increased polymerase activity of a cellular polymerase capable of generating additional endogenous siNA molecules having sequence homology to a chemically-modified siNA molecule comprising (a) introducing nucleotides having  
5 any of Formula I-VII or any combination thereof into a siNA molecule, and (b) assaying the siNA molecule of step (a) under conditions suitable for isolating siNA molecules capable of mediating increased polymerase activity of a cellular polymerase capable of generating additional endogenous siNA molecules having sequence homology to the chemically-modified siNA molecule.

10 In one embodiment, the invention features chemically-modified siNA constructs that mediate RNAi against a HIV in a cell, wherein the chemical modifications do not significantly effect the interaction of siNA with a target RNA molecule, DNA molecule and/or proteins or other factors that are essential for RNAi in a manner that would decrease the efficacy of RNAi mediated by such siNA constructs.

15 In another embodiment, the invention features a method for generating siNA molecules with improved RNAi activity against HIV comprising (a) introducing nucleotides having any of Formula I-VII or any combination thereof into a siNA molecule, and (b) assaying the siNA molecule of step (a) under conditions suitable for isolating siNA molecules having improved RNAi activity.

20 In yet another embodiment, the invention features a method for generating siNA molecules with improved RNAi activity against a HIV target RNA comprising (a) introducing nucleotides having any of Formula I-VII or any combination thereof into a siNA molecule, and (b) assaying the siNA molecule of step (a) under conditions suitable for isolating siNA molecules having improved RNAi activity against the target RNA.

25 In yet another embodiment, the invention features a method for generating siNA molecules with improved RNAi activity against a HIV target DNA comprising (a) introducing nucleotides having any of Formula I-VII or any combination thereof into a siNA molecule, and (b) assaying the siNA molecule of step (a) under conditions suitable for isolating siNA molecules having improved RNAi activity against the target DNA.

In one embodiment, the invention features siNA constructs that mediate RNAi against a HIV, wherein the siNA construct comprises one or more chemical modifications described herein that modulates the cellular uptake of the siNA construct.

In another embodiment, the invention features a method for generating siNA molecules against HIV with improved cellular uptake comprising (a) introducing  
5 nucleotides having any of Formula I-VII or any combination thereof into a siNA molecule, and (b) assaying the siNA molecule of step (a) under conditions suitable for isolating siNA molecules having improved cellular uptake.

In one embodiment, the invention features siNA constructs that mediate RNAi  
10 against a HIV, wherein the siNA construct comprises one or more chemical modifications described herein that increases the bioavailability of the siNA construct, for example, by attaching polymeric conjugates such as polyethyleneglycol or equivalent conjugates that improve the pharmacokinetics of the siNA construct, or by attaching conjugates that target specific tissue types or cell types *in vivo*. Non-limiting examples  
15 of such conjugates are described in Vargeese *et al.*, U.S. Serial No. 10/201,394 incorporated by reference herein.

In one embodiment, the invention features a method for generating siNA molecules of the invention with improved bioavailability, comprising (a) introducing a conjugate into the structure of a siNA molecule, and (b) assaying the siNA molecule of  
20 step (a) under conditions suitable for isolating siNA molecules having improved bioavailability. Such conjugates can include ligands for cellular receptors, such as peptides derived from naturally occurring protein ligands; protein localization sequences, including cellular ZIP code sequences; antibodies; nucleic acid aptamers; vitamins and other co-factors, such as folate and N-acetylgalactosamine; polymers, such as  
25 polyethyleneglycol (PEG); phospholipids; polyamines, such as spermine or spermidine; and others.

In another embodiment, the invention features a method for generating siNA molecules of the invention with improved bioavailability comprising (a) introducing an excipient formulation to a siNA molecule, and (b) assaying the siNA molecule of step (a)  
30 under conditions suitable for isolating siNA molecules having improved bioavailability.

Such excipients include polymers such as cyclodextrins, lipids, cationic lipids, polyamines, phospholipids, and others.

In another embodiment, the invention features a method for generating siNA molecules of the invention with improved bioavailability comprising (a) introducing  
5 nucleotides having any of Formulae I-VII or any combination thereof into a siNA molecule, and (b) assaying the siNA molecule of step (a) under conditions suitable for isolating siNA molecules having improved bioavailability.

In another embodiment, polyethylene glycol (PEG) can be covalently attached to siNA compounds of the present invention. The attached PEG can be any molecular  
10 weight, preferably from about 2,000 to about 50,000 daltons (Da).

The present invention can be used alone or as a component of a kit having at least one of the reagents necessary to carry out the *in vitro* or *in vivo* introduction of RNA to test samples and/or subjects. For example, preferred components of the kit include a siNA molecule of the invention and a vehicle that promotes introduction of the siNA into  
15 cells of interest as described herein (e.g., using lipids and other methods of transfection known in the art, see for example Beigelman *et al.*, US 6,395,713). The kit can be used for target validation, such as in determining gene function and/or activity, or in drug optimization, and in drug discovery (see for example Usman *et al.*, USSN 60/402,996). Such a kit can also include instructions to allow a user of the kit to practice the invention.

20 The term "short interfering nucleic acid", "siNA", "short interfering RNA", "siRNA", "short interfering nucleic acid molecule", "short interfering oligonucleotide molecule", or "chemically-modified short interfering nucleic acid molecule" as used herein refers to any nucleic acid molecule capable of inhibiting or down regulating gene expression or viral replication, for example by mediating RNA interference "RNAi" or  
25 gene silencing in a sequence-specific manner; see for example Bass, 2001, *Nature*, 411, 428-429; Elbashir *et al.*, 2001, *Nature*, 411, 494-498; and Kreutzer *et al.*, International PCT Publication No. WO 00/44895; Zernicka-Goetz *et al.*, International PCT Publication No. WO 01/36646; Fire, International PCT Publication No. WO 99/32619; Plaetinck *et al.*, International PCT Publication No. WO 00/01846; Mello and Fire,  
30 International PCT Publication No. WO 01/29058; Deschamps-Depaillette, International PCT Publication No. WO 99/07409; and Li *et al.*, International PCT Publication No. WO



00/44914; Allshire, 2002, *Science*, 297, 1818-1819; Volpe *et al.*, 2002, *Science*, 297, 1833-1837; Jenuwein, 2002, *Science*, 297, 2215-2218; and Hall *et al.*, 2002, *Science*, 297, 2232-2237; Hutvagner and Zamore, 2002, *Science*, 297, 2056-60; McManus *et al.*, 2002, *RNA*, 8, 842-850; Reinhart *et al.*, 2002, *Gene & Dev.*, 16, 1616-1626; and Reinhart  
5 & Bartel, 2002, *Science*, 297, 1831). Non limiting examples of siNA molecules of the invention are shown in **Figures 4-6**, and **Tables II, III, and IV** herein. For example the siNA can be a double-stranded polynucleotide molecule comprising self-complementary sense and antisense regions, wherein the antisense region comprises nucleotide sequence that is complementary to nucleotide sequence in a target nucleic acid molecule or a  
10 portion thereof and the sense region having nucleotide sequence corresponding to the target nucleic acid sequence or a portion thereof. The siNA can be assembled from two separate oligonucleotides, where one strand is the sense strand and the other is the antisense strand, wherein the antisense and sense strands are self-complementary (i.e. each strand comprises nucleotide sequence that is complementary to nucleotide sequence  
15 in the other strand; such as where the antisense strand and sense strand form a duplex or double stranded structure, for example wherein the double stranded region is about 19 base pairs); the antisense strand comprises nucleotide sequence that is complementary to nucleotide sequence in a target nucleic acid molecule or a portion thereof and the sense strand comprises nucleotide sequence corresponding to the target nucleic acid sequence  
20 or a portion thereof. Alternatively, the siNA is assembled from a single oligonucleotide, where the self-complementary sense and antisense regions of the siNA are linked by means of a nucleic acid based or non-nucleic acid-based linker(s). The siNA can be a polynucleotide with a hairpin secondary structure, having self-complementary sense and antisense regions, wherein the antisense region comprises nucleotide sequence that is  
25 complementary to nucleotide sequence in a separate target nucleic acid molecule or a portion thereof and the sense region having nucleotide sequence corresponding to the target nucleic acid sequence or a portion thereof. The siNA can be a circular single-stranded polynucleotide having two or more loop structures and a stem comprising self-complementary sense and antisense regions, wherein the antisense region comprises  
30 nucleotide sequence that is complementary to nucleotide sequence in a target nucleic acid molecule or a portion thereof and the sense region having nucleotide sequence corresponding to the target nucleic acid sequence or a portion thereof, and wherein the circular polynucleotide can be processed either *in vivo* or *in vitro* to generate an active

siNA molecule capable of mediating RNAi. The siNA can also comprise a single stranded polynucleotide having nucleotide sequence complementary to nucleotide sequence in a target nucleic acid molecule or a portion thereof (for example, where such siNA molecule does not require the presence within the siNA molecule of nucleotide sequence corresponding to the target nucleic acid sequence or a portion thereof), wherein the single stranded polynucleotide can further comprise a terminal phosphate group, such as a 5'-phosphate (see for example Martinez *et al.*, 2002, *Cell.*, 110, 563-574 and Schwarz *et al.*, 2002, *Molecular Cell*, 10, 537-568), or 5',3'-diphosphate. In certain embodiment, the siNA molecule of the invention comprises separate sense and antisense sequences or regions, wherein the sense and antisense regions are covalently linked by nucleotide or non-nucleotide linkers molecules as is known in the art, or are alternately non-covalently linked by ionic interactions, hydrogen bonding, van der waals interactions, hydrophobic interactions, and/or stacking interactions. In certain embodiments, the siNA molecules of the invention comprise nucleotide sequence that is complementary to nucleotide sequence of a target gene. In another embodiment, the siNA molecule of the invention interacts with nucleotide sequence of a target gene in a manner that causes inhibition of expression of the target gene. As used herein, siNA molecules need not be limited to those molecules containing only RNA, but further encompasses chemically-modified nucleotides and non-nucleotides. In certain embodiments, the short interfering nucleic acid molecules of the invention lack 2'-hydroxy (2'-OH) containing nucleotides. Applicant describes in certain embodiments short interfering nucleic acids that do not require the presence of nucleotides having a 2'-hydroxy group for mediating RNAi and as such, short interfering nucleic acid molecules of the invention optionally do not include any ribonucleotides (e.g., nucleotides having a 2'-OH group). Such siNA molecules that do not require the presence of ribonucleotides within the siNA molecule to support RNAi can however have an attached linker or linkers or other attached or associated groups, moieties, or chains containing one or more nucleotides with 2'-OH groups. Optionally, siNA molecules can comprise ribonucleotides at about 5, 10, 20, 30, 40, or 50% of the nucleotide positions. The modified short interfering nucleic acid molecules of the invention can also be referred to as short interfering modified oligonucleotides "siMON." As used herein, the term siNA is meant to be equivalent to other terms used to describe nucleic acid molecules that are capable of mediating sequence specific RNAi, for example short interfering RNA

(siRNA), double-stranded RNA (dsRNA), micro-RNA (miRNA), short hairpin RNA (shRNA), short interfering oligonucleotide, short interfering nucleic acid, short interfering modified oligonucleotide, chemically-modified siRNA, post-transcriptional gene silencing RNA (ptgsRNA), and others. In addition, as used herein, the term RNAi  
5 is meant to be equivalent to other terms used to describe sequence specific RNA interference, such as post transcriptional gene silencing, or epigenetics. For example, siNA molecules of the invention can be used to epigenetically silence genes at both the post-transcriptional level or the pre-transcriptional level. In a non-limiting example, epigenetic regulation of gene expression by siNA molecules of the invention can result  
10 from siNA mediated modification of chromatin structure to alter gene expression (see, for example, Allshire, 2002, *Science*, 297, 1818-1819; Volpe *et al.*, 2002, *Science*, 297, 1833-1837; Jenuwein, 2002, *Science*, 297, 2215-2218; and Hall *et al.*, 2002, *Science*, 297, 2232-2237).

By "modulate" is meant that the expression of the gene, or level of RNA molecule  
15 or equivalent RNA molecules encoding one or more proteins or protein subunits, or activity of one or more proteins or protein subunits is up regulated or down regulated, such that expression, level, or activity is greater than or less than that observed in the absence of the modulator. For example, the term "modulate" can mean "inhibit," but the use of the word "modulate" is not limited to this definition.

20 By "inhibit", "down-regulate", or "reduce", it is meant that the expression of the gene, or level of RNA molecules or equivalent RNA molecules encoding one or more proteins or protein subunits, or activity of one or more proteins or protein subunits, is reduced below that observed in the absence of the nucleic acid molecules (e.g., siNA) of the invention. In one embodiment, inhibition, down-regulation or reduction with an siNA  
25 molecule is below that level observed in the presence of an inactive or attenuated molecule. In another embodiment, inhibition, down-regulation, or reduction with siNA molecules is below that level observed in the presence of, for example, an siNA molecule with scrambled sequence or with mismatches. In another embodiment, inhibition, down-regulation, or reduction of gene expression with a nucleic acid molecule  
30 of the instant invention is greater in the presence of the nucleic acid molecule than in its absence.

By "gene" or "target gene" is meant, a nucleic acid that encodes an RNA, for example, nucleic acid sequences including, but not limited to, structural genes encoding a polypeptide. The target gene can be a gene derived from a cell, an endogenous gene, a transgene, or exogenous genes such as genes of a pathogen, for example a virus, which is present in the cell after infection thereof. The cell containing the target gene can be derived from or contained in any organism, for example a plant, animal, protozoan, virus, bacterium, or fungus. Non-limiting examples of plants include monocots, dicots, or gymnosperms. Non-limiting examples of animals include vertebrates or invertebrates. Non-limiting examples of fungi include molds or yeasts.

By "HIV" as used herein is meant, any virus, protein, peptide, polypeptide, and/or polynucleotide involved in the progression, development, or maintenance of human immunodeficiency virus (HIV) infection and acquired immunodeficiency syndrome (AIDS), including those expressed from a HIV gene or involved in HIV infection (such as polynucleotides referred to by Genbank Accession numbers herein or any other HIV transcript derived from a HIV gene); for example, entire viruses such as HIV-1, HIV-2, FIV-1, SIV-1; viral components such as *nef*, *vif*, *tat*, or *rev* viral gene products; and cellular targets that are involved in HIV infection.

By "HIV protein" is meant, any HIV peptide or protein or a component thereof, wherein the peptide or protein is encoded by a HIV gene or having HIV activity.

By "highly conserved sequence region" is meant, a nucleotide sequence of one or more regions in a target gene does not vary significantly from one generation to the other or from one biological system to the other.

By "sense region" is meant a nucleotide sequence of a siNA molecule having complementarity to an antisense region of the siNA molecule. In addition, the sense region of a siNA molecule can comprise a nucleic acid sequence having homology with a target nucleic acid sequence.

By "antisense region" is meant a nucleotide sequence of a siNA molecule having complementarity to a target nucleic acid sequence. In addition, the antisense region of a siNA molecule can optionally comprise a nucleic acid sequence having complementarity to a sense region of the siNA molecule.

By "target nucleic acid" is meant any nucleic acid sequence whose expression or activity is to be modulated. The target nucleic acid can be DNA or RNA.

By "complementarity" is meant that a nucleic acid can form hydrogen bond(s) with another nucleic acid sequence by either traditional Watson-Crick or other non-traditional types. In reference to the nucleic molecules of the present invention, the binding free energy for a nucleic acid molecule with its complementary sequence is sufficient to allow the relevant function of the nucleic acid to proceed, e.g., RNAi activity. Determination of binding free energies for nucleic acid molecules is well known in the art (see, e.g., Turner *et al.*, 1987, *CSH Symp. Quant. Biol.* LII pp.123-133; Frier *et al.*, 1986, *Proc. Nat. Acad. Sci. USA* 83:9373-9377; Turner *et al.*, 1987, *J. Am. Chem. Soc.* 109:3783-3785). A percent complementarity indicates the percentage of contiguous residues in a nucleic acid molecule that can form hydrogen bonds (e.g., Watson-Crick base pairing) with a second nucleic acid sequence (e.g., 5, 6, 7, 8, 9, 10 out of 10 being 50%, 60%, 70%, 80%, 90%, and 100% complementary). "Perfectly complementary" means that all the contiguous residues of a nucleic acid sequence will hydrogen bond with the same number of contiguous residues in a second nucleic acid sequence.

The siNA molecules of the invention represent a novel therapeutic approach to treat a variety of pathologic indications or other conditions, such as HIV infection or acquired immunodeficiency syndrome (AIDS) and any other diseases or conditions that are related to the levels of HIV in a cell or tissue, alone or in combination with other therapies. The reduction of HIV expression (specifically HIV RNA levels) and thus reduction in the level of the respective protein(s) relieves, to some extent, the symptoms of the disease or condition.

In one embodiment of the present invention, each sequence of a siNA molecule of the invention is independently about 18 to about 24 nucleotides in length, in specific embodiments about 18, 19, 20, 21, 22, 23, or 24 nucleotides in length. In another embodiment, the siNA duplexes of the invention independently comprise about 17 to about 23 base pairs (e.g., about 17, 18, 19, 20, 21, 22 or 23). In yet another embodiment, siNA molecules of the invention comprising hairpin or circular structures are about 35 to about 55 (e.g., about 35, 40, 45, 50 or 55) nucleotides in length, or about 38 to about 44 (e.g., 38, 39, 40, 41, 42, 43 or 44) nucleotides in length and comprising about 16 to about

22 (e.g., about 16, 17, 18, 19, 20, 21 or 22) base pairs. Exemplary siNA molecules of the invention are disclosed herein. Exemplary synthetic siNA molecules of the invention are shown in **Tables III and IV** and/or **Figures 4-5**.

As used herein "cell" is used in its usual biological sense, and does not refer to an entire multicellular organism, e.g., specifically does not refer to a human. The cell can be present in an organism, e.g., birds, plants and mammals such as humans, cows, sheep, apes, monkeys, swine, dogs, and cats. The cell can be prokaryotic (e.g., bacterial cell) or eukaryotic (e.g., mammalian or plant cell). The cell can be of somatic or germ line origin, totipotent or pluripotent, dividing or non-dividing. The cell can also be derived from or can comprise a gamete or embryo, a stem cell, or a fully differentiated cell.

The siNA molecules of the invention are added directly, or can be complexed with cationic lipids, packaged within liposomes, or otherwise delivered to target cells or tissues. The nucleic acid or nucleic acid complexes can be locally administered to relevant tissues *ex vivo*, or *in vivo* through injection, infusion pump or stent, with or without their incorporation in biopolymers. In particular embodiments, the nucleic acid molecules of the invention comprise sequences shown in **Tables II-III** and/or **Figures 4-5**. Examples of such nucleic acid molecules consist essentially of sequences defined in these tables and figures. Furthermore, the chemically modified constructs described in **Table IV** can be applied to any siNA sequence of the invention.

In another aspect, the invention provides mammalian cells containing one or more siNA molecules of this invention. The one or more siNA molecules can independently be targeted to the same or different sites.

By "RNA" is meant a molecule comprising at least one ribonucleotide residue. By "ribonucleotide" is meant a nucleotide with a hydroxyl group at the 2' position of a  $\beta$ -D-ribo-furanose moiety. The terms include double-stranded RNA, single-stranded RNA, isolated RNA such as partially purified RNA, essentially pure RNA, synthetic RNA, recombinantly produced RNA, as well as altered RNA that differs from naturally occurring RNA by the addition, deletion, substitution and/or alteration of one or more nucleotides. Such alterations can include addition of non-nucleotide material, such as to the end(s) of the siNA or internally, for example at one or more nucleotides of the RNA. Nucleotides in the RNA molecules of the instant invention can also comprise non-

standard nucleotides, such as non-naturally occurring nucleotides or chemically synthesized nucleotides or deoxynucleotides. These altered RNAs can be referred to as analogs or analogs of naturally-occurring RNA.

By "subject" is meant an organism, which is a donor or recipient of explanted cells or the cells themselves. "Subject" also refers to an organism to which the nucleic acid molecules of the invention can be administered. In one embodiment, a subject is a mammal or mammalian cells. In another embodiment, a subject is a human or human cells.

The term "phosphorothioate" as used herein refers to an internucleotide linkage having Formula I, wherein Z and/or W comprise a sulfur atom. Hence, the term phosphorothioate refers to both phosphorothioate and phosphorodithioate internucleotide linkages.

The term "universal base" as used herein refers to nucleotide base analogs that form base pairs with each of the natural DNA/RNA bases with little discrimination between them. Non-limiting examples of universal bases include C-phenyl, C-naphthyl and other aromatic derivatives, inosine, azole carboxamides, and nitroazole derivatives such as 3-nitropyrrole, 4-nitroindole, 5-nitroindole, and 6-nitroindole as known in the art (see for example Loakes, 2001, *Nucleic Acids Research*, 29, 2437-2447).

The term "acyclic nucleotide" as used herein refers to any nucleotide having an acyclic ribose sugar, for example where any of the ribose carbons (C1, C2, C3, C4, or C5), are independently or in combination absent from the nucleotide.

The nucleic acid molecules of the instant invention, individually, or in combination or in conjunction with other drugs, can be used to treat diseases or conditions discussed herein (e.g., cancers and other proliferative conditions). For example, to treat a particular disease or condition, the siNA molecules can be administered to a subject or can be administered to other appropriate cells evident to those skilled in the art, individually or in combination with one or more drugs under conditions suitable for the treatment.

In a further embodiment, the siNA molecules can be used in combination with other known treatments to treat conditions or diseases discussed above. For example, the described molecules could be used in combination with one or more known therapeutic

agents to treat a disease or condition. Non-limiting examples of other therapeutic agents that can be readily combined with a siNA molecule of the invention are enzymatic nucleic acid molecules, allosteric nucleic acid molecules, antisense, decoy, or aptamer nucleic acid molecules, antibodies such as monoclonal antibodies, small molecules, and  
5 other organic and/or inorganic compounds including metals, salts and ions.

In one embodiment, the invention features an expression vector comprising a nucleic acid sequence encoding at least one siNA molecule of the invention in a manner which allows expression of the siNA molecule. For example, the vector can contain sequence(s) encoding both strands of a siNA molecule comprising a duplex. The vector  
10 can also contain sequence(s) encoding a single nucleic acid molecule that is self-complementary and thus forms a siNA molecule. Non-limiting examples of such expression vectors are described in Paul *et al.*, 2002, *Nature Biotechnology*, 19, 505; Miyagishi and Taira, 2002, *Nature Biotechnology*, 19, 497; Lee *et al.*, 2002, *Nature Biotechnology*, 19, 500; and Novina *et al.*, 2002, *Nature Medicine*, 8, 681-686.

15 In another embodiment, the invention features a mammalian cell, for example, a human cell, including an expression vector of the invention.

In yet another embodiment, the expression vector of the invention comprises a sequence for a siNA molecule having complementarity to a RNA molecule referred to by a Genbank Accession numbers, for example Genbank Accession Nos. disclosed herein.

20 In one embodiment, an expression vector of the invention comprises a nucleic acid sequence encoding two or more siNA molecules, which can be the same or different.

In another aspect of the invention, siNA molecules that interact with target RNA molecules and down-regulate gene encoding target RNA molecules (for example target RNA molecules referred to by Genbank Accession numbers herein) are expressed from  
25 transcription units inserted into DNA or RNA vectors. The recombinant vectors can be DNA plasmids or viral vectors. siNA expressing viral vectors can be constructed based on, but not limited to, adeno-associated virus, retrovirus, adenovirus, or alphavirus. The recombinant vectors capable of expressing the siNA molecules can be delivered as described herein, and persist in target cells. Alternatively, viral vectors can be used that  
30 provide for transient expression of siNA molecules. Such vectors can be repeatedly



administered as necessary. Once expressed, the siNA molecules bind and down-regulate gene function or expression via RNA interference (RNAi). Delivery of siNA expressing vectors can be systemic, such as by intravenous or intramuscular administration, by administration to target cells ex-planted from a subject followed by reintroduction into the subject, or by any other means that would allow for introduction into the desired target cell.

By "vectors" is meant any nucleic acid- and/or viral-based technique used to deliver a desired nucleic acid.

Other features and advantages of the invention will be apparent from the following description of the preferred embodiments thereof, and from the claims.

### BRIEF DESCRIPTION OF THE DRAWINGS

**Figure 1** shows a non-limiting example of a scheme for the synthesis of siNA molecules. The complementary siNA sequence strands, strand 1 and strand 2, are synthesized in tandem and are connected by a cleavable linkage, such as a nucleotide succinate or abasic succinate, which can be the same or different from the cleavable linker used for solid phase synthesis on a solid support. The synthesis can be either solid phase or solution phase, in the example shown, the synthesis is a solid phase synthesis. The synthesis is performed such that a protecting group, such as a dimethoxytrityl group, remains intact on the terminal nucleotide of the tandem oligonucleotide. Upon cleavage and deprotection of the oligonucleotide, the two siNA strands spontaneously hybridize to form a siNA duplex, which allows the purification of the duplex by utilizing the properties of the terminal protecting group, for example by applying a trityl on purification method wherein only duplexes/oligonucleotides with the terminal protecting group are isolated.

**Figure 2** shows a MALDI-TOV mass spectrum of a purified siNA duplex synthesized by a method of the invention. The two peaks shown correspond to the predicted mass of the separate siNA sequence strands. This result demonstrates that the siNA duplex generated from tandem synthesis can be purified as a single entity using a simple trityl-on purification methodology.

**Figure 3** shows a non-limiting proposed mechanistic representation of target RNA degradation involved in RNAi. Double-stranded RNA (dsRNA), which is generated by RNA-dependent RNA polymerase (RdRP) from foreign single-stranded RNA, for example viral, transposon, or other exogenous RNA, activates the DICER enzyme that in turn generates siNA duplexes. Alternately, synthetic or expressed siNA can be introduced directly into a cell by appropriate means. An active siNA complex forms which recognizes a target RNA, resulting in degradation of the target RNA by the RISC endonuclease complex or in the synthesis of additional RNA by RNA-dependent RNA polymerase (RdRP), which can activate DICER and result in additional siNA molecules, thereby amplifying the RNAi response.

**Figure 4A-F** shows non-limiting examples of chemically-modified siNA constructs of the present invention. In the figure, N stands for any nucleotide (adenosine, guanosine, cytosine, uridine, or optionally thymidine, for example thymidine can be substituted in the overhanging regions designated by parenthesis (N N)). Various modifications are shown for the sense and antisense strands of the siNA constructs.

**Figure 4A:** The sense strand comprises 21 nucleotides having four phosphorothioate 5'- and 3'-terminal internucleotide linkages, wherein the two terminal 3'-nucleotides are optionally base paired and wherein all pyrimidine nucleotides that may be present are 2'-O-methyl or 2'-deoxy-2'-fluoro modified nucleotides except for (N N) nucleotides, which can comprise ribonucleotides, deoxynucleotides, universal bases, or other chemical modifications described herein. The antisense strand comprises 21 nucleotides, optionally having a 3'-terminal glyceryl moiety and wherein the two terminal 3'-nucleotides are optionally complementary to the target RNA sequence, and having one 3'-terminal phosphorothioate internucleotide linkage and four 5'-terminal phosphorothioate internucleotide linkages and wherein all pyrimidine nucleotides that may be present are 2'-deoxy-2'-fluoro modified nucleotides except for (N N) nucleotides, which can comprise ribonucleotides, deoxynucleotides, universal bases, or other chemical modifications described herein.

**Figure 4B:** The sense strand comprises 21 nucleotides wherein the two terminal 3'-nucleotides are optionally base paired and wherein all pyrimidine nucleotides that may be present are 2'-O-methyl or 2'-deoxy-2'-fluoro modified nucleotides except for (N N)

nucleotides, which can comprise ribonucleotides, deoxynucleotides, universal bases, or other chemical modifications described herein. The antisense strand comprises 21 nucleotides, optionally having a 3'-terminal glyceryl moiety and wherein the two terminal 3'-nucleotides are optionally complementary to the target RNA sequence, and  
5 wherein all pyrimidine nucleotides that may be present are 2'-deoxy-2'-fluoro modified nucleotides except for (N N) nucleotides, which can comprise ribonucleotides, deoxynucleotides, universal bases, or other chemical modifications described herein.

**Figure 4C:** The sense strand comprises 21 nucleotides having 5'- and 3'- terminal cap moieties wherein the two terminal 3'-nucleotides are optionally base paired and  
10 wherein all pyrimidine nucleotides that may be present are 2'-O-methyl or 2'-deoxy-2'-fluoro modified nucleotides except for (N N) nucleotides, which can comprise ribonucleotides, deoxynucleotides, universal bases, or other chemical modifications described herein. The antisense strand comprises 21 nucleotides, optionally having a 3'-terminal glyceryl moiety and wherein the two terminal 3'-nucleotides are optionally  
15 complementary to the target RNA sequence, and having one 3'-terminal phosphorothioate internucleotide linkage and wherein all pyrimidine nucleotides that may be present are 2'-deoxy-2'-fluoro modified nucleotides except for (N N) nucleotides, which can comprise ribonucleotides, deoxynucleotides, universal bases, or other chemical modifications described herein.

**Figure 4D:** The sense strand comprises 21 nucleotides having 5'- and 3'- terminal cap moieties wherein the two terminal 3'-nucleotides are optionally base paired and  
20 wherein all pyrimidine nucleotides that may be present are 2'-deoxy-2'-fluoro modified nucleotides except for (N N) nucleotides, which can comprise ribonucleotides, deoxynucleotides, universal bases, or other chemical modifications described herein and  
25 wherein and all purine nucleotides that may be present are 2'-deoxy nucleotides. The antisense strand comprises 21 nucleotides, optionally having a 3'-terminal glyceryl moiety and wherein the two terminal 3'-nucleotides are optionally complementary to the target RNA sequence, and having one 3'-terminal phosphorothioate internucleotide linkage and wherein all pyrimidine nucleotides that may be present are 2'-deoxy-2'-fluoro  
30 modified nucleotides and all purine nucleotides that may be present are 2'-O-methyl modified nucleotides except for (N N) nucleotides, which can comprise ribonucleotides, deoxynucleotides, universal bases, or other chemical modifications described herein.

**Figure 4E:** The sense strand comprises 21 nucleotides having 5'- and 3'- terminal cap moieties wherein the two terminal 3'-nucleotides are optionally base paired and wherein all pyrimidine nucleotides that may be present are 2'-deoxy-2'-fluoro modified nucleotides except for (N N) nucleotides, which can comprise ribonucleotides, deoxynucleotides, universal bases, or other chemical modifications described herein. The antisense strand comprises 21 nucleotides, optionally having a 3'-terminal glyceryl moiety and wherein the two terminal 3'-nucleotides are optionally complementary to the target RNA sequence, and wherein all pyrimidine nucleotides that may be present are 2'-deoxy-2'-fluoro modified nucleotides and all purine nucleotides that may be present are 2'-O-methyl modified nucleotides except for (N N) nucleotides, which can comprise ribonucleotides, deoxynucleotides, universal bases, or other chemical modifications described herein.

**Figure 4F:** The sense strand comprises 21 nucleotides having 5'- and 3'- terminal cap moieties wherein the two terminal 3'-nucleotides are optionally base paired and wherein all pyrimidine nucleotides that may be present are 2'-deoxy-2'-fluoro modified nucleotides except for (N N) nucleotides, which can comprise ribonucleotides, deoxynucleotides, universal bases, or other chemical modifications described herein. The antisense strand comprises 21 nucleotides, optionally having a 3'-terminal glyceryl moiety and wherein the two terminal 3'-nucleotides are optionally complementary to the target RNA sequence, and having one 3'-terminal phosphorothioate internucleotide linkage and wherein all pyrimidine nucleotides that may be present are 2'-deoxy-2'-fluoro modified nucleotides and all purine nucleotides that may be present are 2'-deoxy nucleotides except for (N N) nucleotides, which can comprise ribonucleotides, deoxynucleotides, universal bases, or other chemical modifications described herein. The antisense strand of constructs A-F comprise sequence complementary to any target nucleic acid sequence of the invention.

**Figure 5A-F** shows non-limiting examples of specific chemically-modified siNA sequences of the invention. A-F applies the chemical modifications described in **Figure 4A-F** to a HIV siNA sequence.

**Figure 6** shows non-limiting examples of different siNA constructs of the invention. The examples shown (constructs 1, 2, and 3) have 19 representative base

pairs; however, different embodiments of the invention include any number of base pairs described herein. Bracketed regions represent nucleotide overhangs, for example comprising about 1, 2, 3, or 4 nucleotides in length, preferably about 2 nucleotides. Constructs 1 and 2 can be used independently for RNAi activity. Construct 2 can  
5 comprise a polynucleotide or non-nucleotide linker, which can optionally be designed as a biodegradable linker. In one embodiment, the loop structure shown in construct 2 can comprise a biodegradable linker that results in the formation of construct 1 *in vivo* and/or *in vitro*. In another example, construct 3 can be used to generate construct 2 under the same principle wherein a linker is used to generate the active siNA construct 2 *in vivo*  
10 and/or *in vitro*, which can optionally utilize another biodegradable linker to generate the active siNA construct 1 *in vivo* and/or *in vitro*. As such, the stability and/or activity of the siNA constructs can be modulated based on the design of the siNA construct for use *in vivo* or *in vitro* and/or *in vitro*.

**Figure 7A-C** is a diagrammatic representation of a scheme utilized in generating  
15 an expression cassette to generate siNA hairpin constructs.

**Figure 7A:** A DNA oligomer is synthesized with a 5'-restriction site (R1) sequence followed by a region having sequence identical (sense region of siNA) to a predetermined HIV target sequence, wherein the sense region comprises, for example,  
20 about 19, 20, 21, or 22 nucleotides (N) in length, which is followed by a loop sequence of defined sequence (X), comprising, for example, about 3 to about 10 nucleotides.

**Figure 7B:** The synthetic construct is then extended by DNA polymerase to generate a hairpin structure having self-complementary sequence that will result in a siNA transcript having specificity for a HIV target sequence and having self-complementary sense and antisense regions.

**Figure 7C:** The construct is heated (for example to about 95°C) to linearize the  
25 sequence, thus allowing extension of a complementary second DNA strand using a primer to the 3'-restriction sequence of the first strand. The double-stranded DNA is then inserted into an appropriate vector for expression in cells. The construct can be designed such that a 3'-terminal nucleotide overhang results from the transcription, for example by  
30 engineering restriction sites and/or utilizing a poly-U termination region as described in Paul *et al.*, 2002, *Nature Biotechnology*, 29, 505-508.

**Figure 8A-C** is a diagrammatic representation of a scheme utilized in generating an expression cassette to generate double-stranded siNA constructs.

**Figure 8A:** A DNA oligomer is synthesized with a 5'-restriction (R1) site sequence followed by a region having sequence identical (sense region of siNA) to a predetermined HIV target sequence, wherein the sense region comprises, for example, about 19, 20, 21, or 22 nucleotides (N) in length, and which is followed by a 3'-restriction site (R2) which is adjacent to a loop sequence of defined sequence (X).

**Figure 8B:** The synthetic construct is then extended by DNA polymerase to generate a hairpin structure having self-complementary sequence.

**Figure 8C:** The construct is processed by restriction enzymes specific to R1 and R2 to generate a double-stranded DNA which is then inserted into an appropriate vector for expression in cells. The transcription cassette is designed such that a U6 promoter region flanks each side of the dsDNA which generates the separate sense and antisense strands of the siNA. Poly T termination sequences can be added to the constructs to generate U overhangs in the resulting transcript.

**Figure 9A-E** is a diagrammatic representation of a method used to determine target sites for siNA mediated RNAi within a particular target nucleic acid sequence, such as messenger RNA.

**Figure 9A:** A pool of siNA oligonucleotides are synthesized wherein the antisense region of the siNA constructs has complementarity to target sites across the target nucleic acid sequence, and wherein the sense region comprises sequence complementary to the antisense region of the siNA.

**Figure 9B&C:** (Figure 9B) The sequences are pooled and are inserted into vectors such that (Figure 9C) transfection of a vector into cells results in the expression of the siNA.

**Figure 9D:** Cells are sorted based on phenotypic change that is associated with modulation of the target nucleic acid sequence.

**Figure 9E:** The siNA is isolated from the sorted cells and is sequenced to identify efficacious target sites within the target nucleic acid sequence.

**Figure 10** shows non-limiting examples of different stabilization chemistries (1-10) that can be used, for example, to stabilize the 3'-end of siNA sequences of the invention, including (1) [3'-3']-inverted deoxyribose; (2) deoxyribonucleotide; (3) [5'-3']-3'-deoxyribonucleotide; (4) [5'-3']-ribonucleotide; (5) [5'-3']-3'-O-methyl ribonucleotide; (6) 3'-glyceryl; (7) [3'-5']-3'-deoxyribonucleotide; (8) [3'-3']-deoxyribonucleotide; (9) [5'-2']-deoxyribonucleotide; and (10) [5'-3']-dideoxyribonucleotide. In addition to modified and unmodified backbone chemistries indicated in the figure, these chemistries can be combined with different backbone modifications as described herein, for example, backbone modifications having Formula I. In addition, the 2'-deoxy nucleotide shown 5' to the terminal modifications shown can be another modified or unmodified nucleotide or non-nucleotide described herein, for example modifications having any of Formulae I-VII or any combination thereof.

**Figure 11** shows a non-limiting example of a strategy used to identify chemically modified siNA constructs of the invention that are nuclease resistance while preserving the ability to mediate RNAi activity. Chemical modifications are introduced into the siNA construct based on educated design parameters (e.g. introducing 2'-modifications, base modifications, backbone modifications, terminal cap modifications etc). The modified construct is tested in an appropriate system (e.g. human serum for nuclease resistance, shown, or an animal model for PK/delivery parameters). In parallel, the siNA construct is tested for RNAi activity, for example in a cell culture system such as a luciferase reporter assay). Lead siNA constructs are then identified which possess a particular characteristic while maintaining RNAi activity, and can be further modified and assayed once again. This same approach can be used to identify siNA-conjugate molecules with improved pharmacokinetic profiles, delivery, and RNAi activity.

## DETAILED DESCRIPTION OF THE INVENTION

### Mechanism of action of Nucleic Acid Molecules of the Invention

The discussion that follows discusses the proposed mechanism of RNA interference mediated by short interfering RNA as is presently known, and is not meant to be limiting and is not an admission of prior art. Applicant demonstrates herein that chemically-modified short interfering nucleic acids possess similar or improved capacity to mediate RNAi as do siRNA molecules and are expected to possess improved stability

and activity *in vivo*; therefore, this discussion is not meant to be limiting only to siRNA and can be applied to siNA as a whole. By "improved capacity to mediate RNAi" or "improved RNAi activity" is meant to include RNAi activity measured *in vitro* and/or *in vivo* where the RNAi activity is a reflection of both the ability of the siNA to mediate RNAi and the stability of the siNAs of the invention. In this invention, the product of these activities can be increased *in vitro* and/or *in vivo* compared to an all RNA siRNA or a siNA containing a plurality of ribonucleotides. In some cases, the activity or stability of the siNA molecule can be decreased (i.e., less than ten-fold), but the overall activity of the siNA molecule is enhanced *in vitro* and/or *in vivo*.

RNA interference refers to the process of sequence specific post-transcriptional gene silencing in animals mediated by short interfering RNAs (siRNAs) (Fire *et al.*, 1998, *Nature*, 391, 806). The corresponding process in plants is commonly referred to as post-transcriptional gene silencing or RNA silencing and is also referred to as quelling in fungi. The process of post-transcriptional gene silencing is thought to be an evolutionarily-conserved cellular defense mechanism used to prevent the expression of foreign genes which is commonly shared by diverse flora and phyla (Fire *et al.*, 1999, *Trends Genet.*, 15, 358). Such protection from foreign gene expression may have evolved in response to the production of double-stranded RNAs (dsRNAs) derived from viral infection or the random integration of transposon elements into a host genome via a cellular response that specifically destroys homologous single-stranded RNA or viral genomic RNA. The presence of dsRNA in cells triggers the RNAi response through a mechanism that has yet to be fully characterized. This mechanism appears to be different from the interferon response that results from dsRNA-mediated activation of protein kinase PKR and 2', 5'-oligoadenylate synthetase resulting in non-specific cleavage of mRNA by ribonuclease L.

The presence of long dsRNAs in cells stimulates the activity of a ribonuclease III enzyme referred to as Dicer. Dicer is involved in the processing of the dsRNA into short pieces of dsRNA known as short interfering RNAs (siRNAs) (Berstein *et al.*, 2001, *Nature*, 409, 363). Short interfering RNAs derived from Dicer activity are typically about 21 to about 23 nucleotides in length and comprise about 19 base pair duplexes. Dicer has also been implicated in the excision of 21- and 22-nucleotide small temporal RNAs (stRNAs) from precursor RNA of conserved structure that are implicated in



translational control (Hutvagner *et al.*, 2001, *Science*, 293, 834). The RNAi response also features an endonuclease complex containing a siRNA, commonly referred to as an RNA-induced silencing complex (RISC), which mediates cleavage of single-stranded RNA having sequence homologous to the siRNA. Cleavage of the target RNA takes place in the middle of the region complementary to the guide sequence of the siRNA duplex (Elbashir *et al.*, 2001, *Genes Dev.*, 15, 188). In addition, RNA interference can also involve small RNA (e.g., micro-RNA or miRNA) mediated gene silencing, presumably through cellular mechanisms that regulate chromatin structure and thereby prevent transcription of target gene sequences (see for example Allshire, 2002, *Science*, 297, 1818-1819; Volpe *et al.*, 2002, *Science*, 297, 1833-1837; Jenuwein, 2002, *Science*, 297, 2215-2218; and Hall *et al.*, 2002, *Science*, 297, 2232-2237). As such, siRNA molecules of the invention can be used to mediate gene silencing via interaction with RNA transcripts or alternately by interaction with particular gene sequences, wherein such interaction results in gene silencing either at the transcriptional level or post-transcriptional level.

RNAi has been studied in a variety of systems. Fire *et al.*, 1998, *Nature*, 391, 806, were the first to observe RNAi in *C. elegans*. Wianny and Goetz, 1999, *Nature Cell Biol.*, 2, 70, describe RNAi mediated by dsRNA in mouse embryos. Hammond *et al.*, 2000, *Nature*, 404, 293, describe RNAi in *Drosophila* cells transfected with dsRNA. Elbashir *et al.*, 2001, *Nature*, 411, 494, describe RNAi induced by introduction of duplexes of synthetic 21-nucleotide RNAs in cultured mammalian cells including human embryonic kidney and HeLa cells. Recent work in *Drosophila* embryonic lysates has revealed certain requirements for siRNA length, structure, chemical composition, and sequence that are essential to mediate efficient RNAi activity. These studies have shown that 21 nucleotide siRNA duplexes are most active when containing two 2-nucleotide 3'-terminal nucleotide overhangs. Furthermore, substitution of one or both siRNA strands with 2'-deoxy or 2'-O-methyl nucleotides abolishes RNAi activity, whereas substitution of 3'-terminal siRNA nucleotides with deoxy nucleotides was shown to be tolerated. Mismatch sequences in the center of the siRNA duplex were also shown to abolish RNAi activity. In addition, these studies also indicate that the position of the cleavage site in the target RNA is defined by the 5'-end of the siRNA guide sequence rather than the 3'-end (Elbashir *et al.*, 2001, *EMBO J.*, 20, 6877). Other studies have indicated that a 5'-phosphate on the target-complementary strand of a siRNA duplex is required for siRNA

activity and that ATP is utilized to maintain the 5'-phosphate moiety on the siRNA (Nykanen *et al.*, 2001, *Cell*, 107, 309); however, siRNA molecules lacking a 5'-phosphate are active when introduced exogenously, suggesting that 5'-phosphorylation of siRNA constructs may occur *in vivo*.

## 5 Synthesis of Nucleic acid Molecules

Synthesis of nucleic acids greater than 100 nucleotides in length is difficult using automated methods, and the therapeutic cost of such molecules is prohibitive. In this invention, small nucleic acid motifs ("small" refers to nucleic acid motifs no more than 100 nucleotides in length, preferably no more than 80 nucleotides in length, and most preferably no more than 50 nucleotides in length; *e.g.*, individual siNA oligonucleotide sequences or siNA sequences synthesized in tandem) are preferably used for exogenous delivery. The simple structure of these molecules increases the ability of the nucleic acid to invade targeted regions of protein and/or RNA structure. Exemplary molecules of the instant invention are chemically synthesized, and others can similarly be synthesized.

Oligonucleotides (*e.g.*, certain modified oligonucleotides or portions of oligonucleotides lacking ribonucleotides) are synthesized using protocols known in the art, for example as described in Caruthers *et al.*, 1992, *Methods in Enzymology* 211, 3-19, Thompson *et al.*, International PCT Publication No. WO 99/54459, Wincott *et al.*, 1995, *Nucleic Acids Res.* 23, 2677-2684, Wincott *et al.*, 1997, *Methods Mol. Bio.*, 74, 59, Brennan *et al.*, 1998, *Biotechnol Bioeng.*, 61, 33-45, and Brennan, U.S. Pat. No. 6,001,311. All of these references are incorporated herein by reference. The synthesis of oligonucleotides makes use of common nucleic acid protecting and coupling groups, such as dimethoxytrityl at the 5'-end, and phosphoramidites at the 3'-end. In a non-limiting example, small scale syntheses are conducted on a 394 Applied Biosystems, Inc. synthesizer using a 0.2  $\mu$ mol scale protocol with a 2.5 min coupling step for 2'-O-methylated nucleotides and a 45 sec coupling step for 2'-deoxy nucleotides or 2'-deoxy-2'-fluoro nucleotides. **Table V** outlines the amounts and the contact times of the reagents used in the synthesis cycle. Alternatively, syntheses at the 0.2  $\mu$ mol scale can be performed on a 96-well plate synthesizer, such as the instrument produced by Protogene (Palo Alto, CA) with minimal modification to the cycle. A 33-fold excess (60  $\mu$ L of 0.11 M = 6.6  $\mu$ mol) of 2'-O-methyl phosphoramidite and a 105-fold excess of S-ethyl tetrazole (60  $\mu$ L of 0.25 M = 15  $\mu$ mol) can be used in each coupling cycle of 2'-O-methyl

residues relative to polymer-bound 5'-hydroxyl. A 22-fold excess (40  $\mu$ L of 0.11 M = 4.4  $\mu$ mol) of deoxy phosphoramidite and a 70-fold excess of S-ethyl tetrazole (40  $\mu$ L of 0.25 M = 10  $\mu$ mol) can be used in each coupling cycle of deoxy residues relative to polymer-bound 5'-hydroxyl. Average coupling yields on the 394 Applied Biosystems, Inc. synthesizer, determined by colorimetric quantitation of the trityl fractions, are typically 97.5-99%. Other oligonucleotide synthesis reagents for the 394 Applied Biosystems, Inc. synthesizer include the following: detritylation solution is 3% TCA in methylene chloride (ABI); capping is performed with 16% *N*-methyl imidazole in THF (ABI) and 10% acetic anhydride/10% 2,6-lutidine in THF (ABI); and oxidation solution is 16.9 mM I<sub>2</sub>, 49 mM pyridine, 9% water in THF (PERSEPTIVE™). Burdick & Jackson Synthesis Grade acetonitrile is used directly from the reagent bottle. S-Ethyltetrazole solution (0.25 M in acetonitrile) is made up from the solid obtained from American International Chemical, Inc. Alternately, for the introduction of phosphorothioate linkages, Beaucage reagent (3H-1,2-Benzodithiol-3-one 1,1-dioxide, 0.05 M in acetonitrile) is used.

Deprotection of the DNA-based oligonucleotides is performed as follows: the polymer-bound trityl-on oligoribonucleotide is transferred to a 4 mL glass screw top vial and suspended in a solution of 40% aq. methylamine (1 mL) at 65 °C for 10 min. After cooling to -20 °C, the supernatant is removed from the polymer support. The support is washed three times with 1.0 mL of EtOH:MeCN:H<sub>2</sub>O/3:1:1, vortexed and the supernatant is then added to the first supernatant. The combined supernatants, containing the oligoribonucleotide, are dried to a white powder.

The method of synthesis used for RNA including certain siNA molecules of the invention follows the procedure as described in Usman *et al.*, 1987, *J. Am. Chem. Soc.*, 109, 7845; Scaringe *et al.*, 1990, *Nucleic Acids Res.*, 18, 5433; and Wincott *et al.*, 1995, *Nucleic Acids Res.* 23, 2677-2684 Wincott *et al.*, 1997, *Methods Mol. Bio.*, 74, 59, and makes use of common nucleic acid protecting and coupling groups, such as dimethoxytrityl at the 5'-end, and phosphoramidites at the 3'-end. In a non-limiting example, small scale syntheses are conducted on a 394 Applied Biosystems, Inc. synthesizer using a 0.2  $\mu$ mol scale protocol with a 7.5 min coupling step for alkylsilyl protected nucleotides and a 2.5 min coupling step for 2'-O-methylated nucleotides. Table V outlines the amounts and the contact times of the reagents used in the synthesis

cycle. Alternatively, syntheses at the 0.2  $\mu\text{mol}$  scale can be done on a 96-well plate synthesizer, such as the instrument produced by Protogene (Palo Alto, CA) with minimal modification to the cycle. A 33-fold excess (60  $\mu\text{L}$  of 0.11 M = 6.6  $\mu\text{mol}$ ) of 2'-O-methyl phosphoramidite and a 75-fold excess of S-ethyl tetrazole (60  $\mu\text{L}$  of 0.25 M = 15  $\mu\text{mol}$ ) can be used in each coupling cycle of 2'-O-methyl residues relative to polymer-bound 5'-hydroxyl. A 66-fold excess (120  $\mu\text{L}$  of 0.11 M = 13.2  $\mu\text{mol}$ ) of alkylsilyl (ribo) protected phosphoramidite and a 150-fold excess of S-ethyl tetrazole (120  $\mu\text{L}$  of 0.25 M = 30  $\mu\text{mol}$ ) can be used in each coupling cycle of ribo residues relative to polymer-bound 5'-hydroxyl. Average coupling yields on the 394 Applied Biosystems, Inc. synthesizer, determined by colorimetric quantitation of the trityl fractions, are typically 97.5-99%. Other oligonucleotide synthesis reagents for the 394 Applied Biosystems, Inc. synthesizer include the following: detritylation solution is 3% TCA in methylene chloride (ABI); capping is performed with 16% *N*-methyl imidazole in THF (ABI) and 10% acetic anhydride/10% 2,6-lutidine in THF (ABI); oxidation solution is 16.9 mM  $\text{I}_2$ , 49 mM pyridine, 9% water in THF (PERSEPTIVE™). Burdick & Jackson Synthesis Grade acetonitrile is used directly from the reagent bottle. S-Ethyltetrazole solution (0.25 M in acetonitrile) is made up from the solid obtained from American International Chemical, Inc. Alternately, for the introduction of phosphorothioate linkages, Beaucage reagent (3H-1,2-Benzodithiol-3-one 1,1-dioxide 0.05 M in acetonitrile) is used.

Deprotection of the RNA is performed using either a two-pot or one-pot protocol. For the two-pot protocol, the polymer-bound trityl-on oligoribonucleotide is transferred to a 4 mL glass screw top vial and suspended in a solution of 40% aq. methylamine (1 mL) at 65 °C for 10 min. After cooling to -20 °C, the supernatant is removed from the polymer support. The support is washed three times with 1.0 mL of EtOH:MeCN:H<sub>2</sub>O/3:1:1, vortexed and the supernatant is then added to the first supernatant. The combined supernatants, containing the oligoribonucleotide, are dried to a white powder. The base deprotected oligoribonucleotide is resuspended in anhydrous TEA/HF/NMP solution (300  $\mu\text{L}$  of a solution of 1.5 mL *N*-methylpyrrolidinone, 750  $\mu\text{L}$  TEA and 1 mL TEA•3HF to provide a 1.4 M HF concentration) and heated to 65 °C. After 1.5 h, the oligomer is quenched with 1.5 M  $\text{NH}_4\text{HCO}_3$ .

Alternatively, for the one-pot protocol, the polymer-bound trityl-on oligoribonucleotide is transferred to a 4 mL glass screw top vial and suspended in a

solution of 33% ethanolic methylamine/DMSO: 1/1 (0.8 mL) at 65 °C for 15 min. The vial is brought to rt. TEA•3HF (0.1 mL) is added and the vial is heated at 65 °C for 15 min. The sample is cooled at -20 °C and then quenched with 1.5 M NH<sub>4</sub>HCO<sub>3</sub>.

For purification of the trityl-on oligomers, the quenched NH<sub>4</sub>HCO<sub>3</sub> solution is  
5 loaded onto a C-18 containing cartridge that had been prewashed with acetonitrile followed by 50 mM TEAA. After washing the loaded cartridge with water, the RNA is detritylated with 0.5% TFA for 13 min. The cartridge is then washed again with water, salt exchanged with 1 M NaCl and washed with water again. The oligonucleotide is then eluted with 30% acetonitrile.

10 The average stepwise coupling yields are typically >98% (Wincott *et al.*, 1995 *Nucleic Acids Res.* 23, 2677-2684). Those of ordinary skill in the art will recognize that the scale of synthesis can be adapted to be larger or smaller than the example described above including but not limited to 96-well format.

Alternatively, the nucleic acid molecules of the present invention can be  
15 synthesized separately and joined together post-synthetically, for example, by ligation (Moore *et al.*, 1992, *Science* 256, 9923; Draper *et al.*, International PCT publication No. WO 93/23569; Shabarova *et al.*, 1991, *Nucleic Acids Research* 19, 4247; Bellon *et al.*, 1997, *Nucleosides & Nucleotides*, 16, 951; Bellon *et al.*, 1997, *Bioconjugate Chem.* 8, 204), or by hybridization following synthesis and/or deprotection.

20 The siNA molecules of the invention can also be synthesized via a tandem synthesis methodology as described in Example 1 herein, wherein both siNA strands are synthesized as a single contiguous oligonucleotide fragment or strand separated by a cleavable linker which is subsequently cleaved to provide separate siNA fragments or strands that hybridize and permit purification of the siNA duplex. The linker can be a  
25 polynucleotide linker or a non-nucleotide linker. The tandem synthesis of siNA as described herein can be readily adapted to both multiwell/multiplate synthesis platforms such as 96 well or similarly larger multi-well platforms. The tandem synthesis of siNA as described herein can also be readily adapted to large scale synthesis platforms employing batch reactors, synthesis columns and the like.

A siNA molecule can also be assembled from two distinct nucleic acid strands or fragments wherein one fragment includes the sense region and the second fragment includes the antisense region of the RNA molecule.

The nucleic acid molecules of the present invention can be modified extensively to enhance stability by modification with nuclease resistant groups, for example, 2'-amino, 2'-C-allyl, 2'-fluoro, 2'-O-methyl, 2'-H (for a review see Usman and Cedergren, 1992, *TIBS* 17, 34; Usman *et al.*, 1994, *Nucleic Acids Symp. Ser.* 31, 163). siNA constructs can be purified by gel electrophoresis using general methods or can be purified by high pressure liquid chromatography (HPLC; see Wincott *et al.*, *supra*, the totality of which is hereby incorporated herein by reference) and re-suspended in water.

In another aspect of the invention, siNA molecules of the invention are expressed from transcription units inserted into DNA or RNA vectors. The recombinant vectors can be DNA plasmids or viral vectors. siNA expressing viral vectors can be constructed based on, but not limited to, adeno-associated virus, retrovirus, adenovirus, or alphavirus. The recombinant vectors capable of expressing the siNA molecules can be delivered as described herein, and persist in target cells. Alternatively, viral vectors can be used that provide for transient expression of siNA molecules.

#### Optimizing Activity of the nucleic acid molecule of the invention.

Chemically synthesizing nucleic acid molecules with modifications (base, sugar and/or phosphate) can prevent their degradation by serum ribonucleases, which can increase their potency (see *e.g.*, Eckstein *et al.*, International Publication No. WO 92/07065; Perrault *et al.*, 1990 *Nature* 344, 565; Pieken *et al.*, 1991, *Science* 253, 314; Usman and Cedergren, 1992, *Trends in Biochem. Sci.* 17, 334; Usman *et al.*, International Publication No. WO 93/15187; and Rossi *et al.*, International Publication No. WO 91/03162; Sproat, U.S. Pat. No. 5,334,711; Gold *et al.*, U.S. Pat. No. 6,300,074; and Burgin *et al.*, *supra*; all of which are incorporated by reference herein). All of the above references describe various chemical modifications that can be made to the base, phosphate and/or sugar moieties of the nucleic acid molecules described herein. Modifications that enhance their efficacy in cells, and removal of bases from nucleic acid molecules to shorten oligonucleotide synthesis times and reduce chemical requirements are desired.

There are several examples in the art describing sugar, base and phosphate modifications that can be introduced into nucleic acid molecules with significant enhancement in their nuclease stability and efficacy. For example, oligonucleotides are modified to enhance stability and/or enhance biological activity by modification with

5 nuclease resistant groups, for example, 2'-amino, 2'-C-allyl, 2'-fluoro, 2'-O-methyl, 2'-O-allyl, 2'-H, nucleotide base modifications (for a review see Usman and Cedergren, 1992, *TIBS*, 17, 34; Usman *et al.*, 1994, *Nucleic Acids Symp. Ser.* 31, 163; Burgin *et al.*, 1996, *Biochemistry*, 35, 14090). Sugar modification of nucleic acid molecules have been extensively described in the art (see Eckstein *et al.*, *International Publication* PCT No.

10 WO 92/07065; Perrault *et al.* *Nature*, 1990, 344, 565-568; Pieken *et al.* *Science*, 1991, 253, 314-317; Usman and Cedergren, *Trends in Biochem. Sci.*, 1992, 17, 334-339; Usman *et al.* *International Publication* PCT No. WO 93/15187; Sproat, U.S. Pat. No. 5,334,711 and Beigelman *et al.*, 1995, *J. Biol. Chem.*, 270, 25702; Beigelman *et al.*, *International PCT publication* No. WO 97/26270; Beigelman *et al.*, U.S. Pat. No.

15 5,716,824; Usman *et al.*, U.S. Pat. No. 5,627,053; Woolf *et al.*, *International PCT Publication* No. WO 98/13526; Thompson *et al.*, USSN 60/082,404 which was filed on April 20, 1998; Karpeisky *et al.*, 1998, *Tetrahedron Lett.*, 39, 1131; Earnshaw and Gait, 1998, *Biopolymers (Nucleic Acid Sciences)*, 48, 39-55; Verma and Eckstein, 1998, *Annu. Rev. Biochem.*, 67, 99-134; and Burlina *et al.*, 1997, *Bioorg. Med. Chem.*, 5, 1999-2010;

20 all of the references are hereby incorporated in their totality by reference herein). Such publications describe general methods and strategies to determine the location of incorporation of sugar, base and/or phosphate modifications and the like into nucleic acid molecules without modulating catalysis, and are incorporated by reference herein. In view of such teachings, similar modifications can be used as described herein to modify

25 the siNA nucleic acid molecules of the instant invention so long as the ability of siNA to promote RNAi in cells is not significantly inhibited.

While chemical modification of oligonucleotide internucleotide linkages with phosphorothioate, phosphorodithioate, and/or 5'-methylphosphonate linkages improves stability, excessive modifications can cause some toxicity or decreased activity.

30 Therefore, when designing nucleic acid molecules, the amount of these internucleotide linkages should be minimized. The reduction in the concentration of these linkages should lower toxicity, resulting in increased efficacy and higher specificity of these molecules.

Short interfering nucleic acid (siNA) molecules having chemical modifications that maintain or enhance activity are provided. Such a nucleic acid is also generally more resistant to nucleases than an unmodified nucleic acid. Accordingly, the *in vitro* and/or *in vivo* activity should not be significantly lowered. In cases in which modulation is the goal, therapeutic nucleic acid molecules delivered exogenously should optimally be stable within cells until translation of the target RNA has been modulated long enough to reduce the levels of the undesirable protein. This period of time varies between hours to days depending upon the disease state. Improvements in the chemical synthesis of RNA and DNA (Wincott *et al.*, 1995, *Nucleic Acids Res.* 23, 2677; Caruthers *et al.*, 1992, *Methods in Enzymology* 211,3-19 (incorporated by reference herein)) have expanded the ability to modify nucleic acid molecules by introducing nucleotide modifications to enhance their nuclease stability, as described above.

In one embodiment, nucleic acid molecules of the invention include one or more (*e.g.*, about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more) G-clamp nucleotides. A G-clamp nucleotide is a modified cytosine analog wherein the modifications confer the ability to hydrogen bond both Watson-Crick and Hoogsteen faces of a complementary guanine within a duplex, see for example Lin and Matteucci, 1998, *J. Am. Chem. Soc.*, 120, 8531-8532. A single G-clamp analog substitution within an oligonucleotide can result in substantially enhanced helical thermal stability and mismatch discrimination when hybridized to complementary oligonucleotides. The inclusion of such nucleotides in nucleic acid molecules of the invention results in both enhanced affinity and specificity to nucleic acid targets, complementary sequences, or template strands. In another embodiment, nucleic acid molecules of the invention include one or more (*e.g.*, about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more) LNA "locked nucleic acid" nucleotides such as a 2', 4'-C methylene bicyclo nucleotide (see for example Wengel *et al.*, International PCT Publication No. WO 00/66604 and WO 99/14226).

In another embodiment, the invention features conjugates and/or complexes of siNA molecules of the invention. Such conjugates and/or complexes can be used to facilitate delivery of siNA molecules into a biological system, such as a cell. The conjugates and complexes provided by the instant invention can impart therapeutic activity by transferring therapeutic compounds across cellular membranes, altering the pharmacokinetics, and/or modulating the localization of nucleic acid molecules of the



invention. The present invention encompasses the design and synthesis of novel conjugates and complexes for the delivery of molecules, including, but not limited to, small molecules, lipids, phospholipids, nucleosides, nucleotides, nucleic acids, antibodies, toxins, negatively charged polymers and other polymers, for example  
5 proteins, peptides, hormones, carbohydrates, polyethylene glycols, or polyamines, across cellular membranes. In general, the transporters described are designed to be used either individually or as part of a multi-component system, with or without degradable linkers. These compounds are expected to improve delivery and/or localization of nucleic acid molecules of the invention into a number of cell types originating from different tissues,  
10 in the presence or absence of serum (see Sullenger and Cech, U.S. Pat. No. 5,854,038). Conjugates of the molecules described herein can be attached to biologically active molecules via linkers that are biodegradable, such as biodegradable nucleic acid linker molecules.

The term "biodegradable linker" as used herein, refers to a nucleic acid or non-  
15 nucleic acid linker molecule that is designed as a biodegradable linker to connect one molecule to another molecule, for example, a biologically active molecule to a siNA molecule of the invention or the sense and antisense strands of a siNA molecule of the invention. The biodegradable linker is designed such that its stability can be modulated for a particular purpose, such as delivery to a particular tissue or cell type. The stability  
20 of a nucleic acid-based biodegradable linker molecule can be modulated by using various chemistries, for example combinations of ribonucleotides, deoxyribonucleotides, and chemically-modified nucleotides, such as 2'-O-methyl, 2'-fluoro, 2'-amino, 2'-O-amino, 2'-C-allyl, 2'-O-allyl, and other 2'-modified or base modified nucleotides. The biodegradable nucleic acid linker molecule can be a dimer, trimer, tetramer or longer  
25 nucleic acid molecule, for example, an oligonucleotide of about 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 nucleotides in length, or can comprise a single nucleotide with a phosphorus-based linkage, for example, a phosphoramidate or phosphodiester linkage. The biodegradable nucleic acid linker molecule can also comprise nucleic acid backbone, nucleic acid sugar, or nucleic acid base modifications.

30 The term "biodegradable" as used herein, refers to degradation in a biological system, for example enzymatic degradation or chemical degradation.

The term "biologically active molecule" as used herein, refers to compounds or molecules that are capable of eliciting or modifying a biological response in a system. Non-limiting examples of biologically active siNA molecules either alone or in combination with other molecules contemplated by the instant invention include

5 therapeutically active molecules such as antibodies, hormones, antivirals, peptides, proteins, chemotherapeutics, small molecules, vitamins, co-factors, nucleosides, nucleotides, oligonucleotides, enzymatic nucleic acids, antisense nucleic acids, triplex forming oligonucleotides, 2,5-A chimeras, siNA, dsRNA, allozymes, aptamers, decoys and analogs thereof. Biologically active molecules of the invention also include

10 molecules capable of modulating the pharmacokinetics and/or pharmacodynamics of other biologically active molecules, for example, lipids and polymers such as polyamines, polyamides, polyethylene glycol and other polyethers.

The term "phospholipid" as used herein, refers to a hydrophobic molecule comprising at least one phosphorus group. For example, a phospholipid can comprise a

15 phosphorus-containing group and saturated or unsaturated alkyl group, optionally substituted with OH, COOH, oxo, amine, or substituted or unsubstituted aryl groups.

Therapeutic nucleic acid molecules (*e.g.*, siNA molecules) delivered exogenously optimally are stable within cells until reverse transcription of the RNA has been modulated long enough to reduce the levels of the RNA transcript. The nucleic acid

20 molecules are resistant to nucleases in order to function as effective intracellular therapeutic agents. Improvements in the chemical synthesis of nucleic acid molecules described in the instant invention and in the art have expanded the ability to modify nucleic acid molecules by introducing nucleotide modifications to enhance their nuclease stability as described above.

25 In yet another embodiment, siNA molecules having chemical modifications that maintain or enhance enzymatic activity of proteins involved in RNAi are provided. Such nucleic acids are also generally more resistant to nucleases than unmodified nucleic acids. Thus, *in vitro* and/or *in vivo* the activity should not be significantly lowered.

Use of the nucleic acid-based molecules of the invention will lead to better

30 treatment of the disease progression by affording the possibility of combination therapies (*e.g.*, multiple siNA molecules targeted to different genes; nucleic acid molecules

coupled with known small molecule modulators; or intermittent treatment with combinations of molecules, including different motifs and/or other chemical or biological molecules). The treatment of subjects with siNA molecules can also include combinations of different types of nucleic acid molecules, such as enzymatic nucleic acid molecules (ribozymes), allozymes, antisense, 2,5-A oligoadenylate, decoys, and aptamers.

In another aspect a siNA molecule of the invention comprises one or more 5' and/or a 3'-cap structure, for example on only the sense siNA strand, the antisense siNA strand, or both siNA strands.

By "cap structure" is meant chemical modifications, which have been incorporated at either terminus of the oligonucleotide (see, for example, Adamic *et al.*, U.S. Pat. No. 5,998,203, incorporated by reference herein). These terminal modifications protect the nucleic acid molecule from exonuclease degradation, and may help in delivery and/or localization within a cell. The cap may be present at the 5'-terminus (5'-cap) or at the 3'-terminal (3'-cap) or may be present on both termini. In non-limiting examples, the 5'-cap is selected from the group consisting of glyceryl, inverted deoxy abasic residue (moiety); 4',5'-methylene nucleotide; 1-(beta-D-erythrofuransyl) nucleotide, 4'-thio nucleotide; carbocyclic nucleotide; 1,5-anhydrohexitol nucleotide; L-nucleotides; alpha-nucleotides; modified base nucleotide; phosphorodithioate linkage; *threo*-pentofuransyl nucleotide; acyclic 3',4'-seco nucleotide; acyclic 3,4-dihydroxybutyl nucleotide; acyclic 3,5-dihydroxypentyl nucleotide, 3'-3'-inverted nucleotide moiety, 3'-3'-inverted abasic moiety; 3'-2'-inverted nucleotide moiety; 3'-2'-inverted abasic moiety; 1,4-butanediol phosphate; 3'-phosphoramidate; hexylphosphate; aminohexyl phosphate; 3'-phosphate; 3'-phosphorothioate; phosphorodithioate; or bridging or non-bridging methylphosphonate moiety.

In non-limiting examples, the 3'-cap is selected from the group consisting of glyceryl, inverted deoxy abasic residue (moiety), 4',5'-methylene nucleotide; 1-(beta-D-erythrofuransyl) nucleotide; 4'-thio nucleotide, carbocyclic nucleotide; 5'-amino-alkyl phosphate; 1,3-diamino-2-propyl phosphate; 3-aminopropyl phosphate; 6-aminohexyl phosphate; 1,2-aminododecyl phosphate; hydroxypropyl phosphate; 1,5-anhydrohexitol nucleotide; L-nucleotide; alpha-nucleotide; modified base nucleotide;

phosphorodithioate; *threo*-pentofuranosyl nucleotide; acyclic 3',4'-seco nucleotide; 3,4-dihydroxybutyl nucleotide; 3,5-dihydroxypentyl nucleotide, 5'-5'-inverted nucleotide moiety; 5'-5'-inverted abasic moiety; 5'-phosphoramidate; 5'-phosphorothioate; 1,4-butanediol phosphate; 5'-amino; bridging and/or non-bridging 5'-phosphoramidate, 5 phosphorothioate and/or phosphorodithioate, bridging or non bridging methylphosphonate and 5'-mercapto moieties (for more details see Beaucage and Iyer, 1993, *Tetrahedron* 49, 1925; incorporated by reference herein).

By the term "non-nucleotide" is meant any group or compound which can be incorporated into a nucleic acid chain in the place of one or more nucleotide units, including either sugar and/or phosphate substitutions, and allows the remaining bases to exhibit their enzymatic activity. The group or compound is abasic in that it does not contain a commonly recognized nucleotide base, such as adenosine, guanine, cytosine, uracil or thymine and therefore lacks a base at the 1'-position.

An "alkyl" group refers to a saturated aliphatic hydrocarbon, including straight-chain, branched-chain, and cyclic alkyl groups. Preferably, the alkyl group has 1 to 12 carbons. More preferably, it is a lower alkyl of from 1 to 7 carbons, more preferably 1 to 4 carbons. The alkyl group can be substituted or unsubstituted. When substituted the substituted group(s) is preferably, hydroxyl, cyano, alkoxy, =O, =S, NO<sub>2</sub> or N(CH<sub>3</sub>)<sub>2</sub>, amino, or SH. The term also includes alkenyl groups that are unsaturated hydrocarbon groups containing at least one carbon-carbon double bond, including straight-chain, branched-chain, and cyclic groups. Preferably, the alkenyl group has 1 to 12 carbons. More preferably, it is a lower alkenyl of from 1 to 7 carbons, more preferably 1 to 4 carbons. The alkenyl group may be substituted or unsubstituted. When substituted the substituted group(s) is preferably, hydroxyl, cyano, alkoxy, =O, =S, NO<sub>2</sub>, halogen, N(CH<sub>3</sub>)<sub>2</sub>, amino, or SH. The term "alkyl" also includes alkynyl groups that have an unsaturated hydrocarbon group containing at least one carbon-carbon triple bond, including straight-chain, branched-chain, and cyclic groups. Preferably, the alkynyl group has 1 to 12 carbons. More preferably, it is a lower alkynyl of from 1 to 7 carbons, more preferably 1 to 4 carbons. The alkynyl group may be substituted or unsubstituted. When substituted the substituted group(s) is preferably, hydroxyl, cyano, alkoxy, =O, =S, NO<sub>2</sub> or N(CH<sub>3</sub>)<sub>2</sub>, amino or SH.

Such alkyl groups can also include aryl, alkylaryl, carbocyclic aryl, heterocyclic aryl, amide and ester groups. An "aryl" group refers to an aromatic group that has at least one ring having a conjugated pi electron system and includes carbocyclic aryl, heterocyclic aryl and biaryl groups, all of which may be optionally substituted. The preferred substituent(s) of aryl groups are halogen, trihalomethyl, hydroxyl, SH, OH, cyano, alkoxy, alkyl, alkenyl, alkynyl, and amino groups. An "alkylaryl" group refers to an alkyl group (as described above) covalently joined to an aryl group (as described above). Carbocyclic aryl groups are groups wherein the ring atoms on the aromatic ring are all carbon atoms. The carbon atoms are optionally substituted. Heterocyclic aryl groups are groups having from 1 to 3 heteroatoms as ring atoms in the aromatic ring and the remainder of the ring atoms are carbon atoms. Suitable heteroatoms include oxygen, sulfur, and nitrogen, and include furanyl, thienyl, pyridyl, pyrrolyl, N-lower alkyl pyrrolo, pyrimidyl, pyrazinyl, imidazolyl and the like, all optionally substituted. An "amide" refers to an -C(O)-NH-R, where R is either alkyl, aryl, alkylaryl or hydrogen. An "ester" refers to an -C(O)-OR', where R is either alkyl, aryl, alkylaryl or hydrogen.

By "nucleotide" as used herein is as recognized in the art to include natural bases (standard), and modified bases well known in the art. Such bases are generally located at the 1' position of a nucleotide sugar moiety. Nucleotides generally comprise a base, sugar and a phosphate group. The nucleotides can be unmodified or modified at the sugar, phosphate and/or base moiety, (also referred to interchangeably as nucleotide analogs, modified nucleotides, non-natural nucleotides, non-standard nucleotides and other; see, for example, Usman and McSwiggen, *supra*; Eckstein *et al.*, International PCT Publication No. WO 92/07065; Usman *et al.*, International PCT Publication No. WO 93/15187; Uhlman & Peyman, *supra*, all are hereby incorporated by reference herein). There are several examples of modified nucleic acid bases known in the art as summarized by Limbach *et al.*, 1994, *Nucleic Acids Res.* 22, 2183. Some of the non-limiting examples of base modifications that can be introduced into nucleic acid molecules include, inosine, purine, pyridin-4-one, pyridin-2-one, phenyl, pseudouracil, 2, 4, 6-trimethoxy benzene, 3-methyl uracil, dihydrouridine, naphthyl, aminophenyl, 5-alkylcytidines (*e.g.*, 5-methylcytidine), 5-alkyluridines (*e.g.*, ribothymidine), 5-halouridine (*e.g.*, 5-bromouridine) or 6-azapyrimidines or 6-alkylpyrimidines (*e.g.* 6-methyluridine), propyne, and others (Burgin *et al.*, 1996, *Biochemistry*, 35, 14090;

Uhlman & Peyman, *supra*). By "modified bases" in this aspect is meant nucleotide bases other than adenine, guanine, cytosine and uracil at 1' position or their equivalents.

In one embodiment, the invention features modified siNA molecules, with phosphate backbone modifications comprising one or more phosphorothioate, phosphorodithioate, methylphosphonate, phosphotriester, morpholino, amidate, carbamate, carboxymethyl, acetamidate, polyamide, sulfonate, sulfonamide, sulfamate, formacetal, thioformacetal, and/or alkylsilyl, substitutions. For a review of oligonucleotide backbone modifications, see Hunziker and Leumann, 1995, *Nucleic Acid Analogues: Synthesis and Properties*, in *Modern Synthetic Methods*, VCH, 331-417, and Mesmaeker *et al.*, 1994, *Novel Backbone Replacements for Oligonucleotides*, in *Carbohydrate Modifications in Antisense Research*, ACS, 24-39.

By "abasic" is meant sugar moieties lacking a base or having other chemical groups in place of a base at the 1' position, see for example Adamic *et al.*, U.S. Pat. No. 5,998,203.

By "unmodified nucleoside" is meant one of the bases adenine, cytosine, guanine, thymine, or uracil joined to the 1' carbon of  $\beta$ -D-ribo-furanose.

By "modified nucleoside" is meant any nucleotide base which contains a modification in the chemical structure of an unmodified nucleotide base, sugar and/or phosphate. Non-limiting examples of modified nucleotides are shown by Formulae I-VII and/or other modifications described herein.

In connection with 2'-modified nucleotides as described for the present invention, by "amino" is meant 2'-NH<sub>2</sub> or 2'-O- NH<sub>2</sub>, which can be modified or unmodified. Such modified groups are described, for example, in Eckstein *et al.*, U.S. Pat. No. 5,672,695 and Matulic-Adamic *et al.*, U.S. Pat. No. 6,248,878, which are both incorporated by reference in their entireties.

Various modifications to nucleic acid siNA structure can be made to enhance the utility of these molecules. Such modifications will enhance shelf-life, half-life *in vitro*, stability, and ease of introduction of such oligonucleotides to the target site, *e.g.*, to enhance penetration of cellular membranes, and confer the ability to recognize and bind to targeted cells.

Administration of Nucleic Acid Molecules

A siRNA molecule of the invention can be adapted for use to treat, for example conditions related to HIV infection and/or AIDS, alone or in combination with other therapies. For example, a siNA molecule can comprise a delivery vehicle, including liposomes, for administration to a subject, carriers and diluents and their salts, and/or can be present in pharmaceutically acceptable formulations. Methods for the delivery of nucleic acid molecules are described in Akhtar *et al.*, 1992, *Trends Cell Bio.*, 2, 139; *Delivery Strategies for Antisense Oligonucleotide Therapeutics*, ed. Akhtar, 1995, Maurer *et al.*, 1999, *Mol. Membr. Biol.*, 16, 129-140; Hofland and Huang, 1999, *Handb. Exp. Pharmacol.*, 137, 165-192; and Lee *et al.*, 2000, *ACS Symp. Ser.*, 752, 184-192, all of which are incorporated herein by reference. Beigelman *et al.*, U.S. Pat. No. 6,395,713 and Sullivan *et al.*, PCT WO 94/02595 further describe the general methods for delivery of nucleic acid molecules. These protocols can be utilized for the delivery of virtually any nucleic acid molecule. Nucleic acid molecules can be administered to cells by a variety of methods known to those of skill in the art, including, but not restricted to, encapsulation in liposomes, by iontophoresis, or by incorporation into other vehicles, such as hydrogels, cyclodextrins (see for example Gonzalez *et al.*, 1999, *Bioconjugate Chem.*, 10, 1068-1074), biodegradable nanocapsules, and bioadhesive microspheres, or by proteinaceous vectors (O'Hare and Normand, International PCT Publication No. WO 00/53722). Alternatively, the nucleic acid/vehicle combination is locally delivered by direct injection or by use of an infusion pump. Direct injection of the nucleic acid molecules of the invention, whether subcutaneous, intramuscular, or intradermal, can take place using standard needle and syringe methodologies, or by needle-free technologies such as those described in Conry *et al.*, 1999, *Clin. Cancer Res.*, 5, 2330-2337 and Barry *et al.*, International PCT Publication No. WO 99/31262. The molecules of the instant invention can be used as pharmaceutical agents. Pharmaceutical agents prevent, modulate the occurrence, or treat (alleviate a symptom to some extent, preferably all of the symptoms) of a disease state in a subject.

Thus, the invention features a pharmaceutical composition comprising one or more nucleic acid(s) of the invention in an acceptable carrier, such as a stabilizer, buffer, and the like. The polynucleotides of the invention can be administered (*e.g.*, RNA, DNA or protein) and introduced into a subject by any standard means, with or without stabilizers,

5 buffers, and the like, to form a pharmaceutical composition. When it is desired to use a liposome delivery mechanism, standard protocols for formation of liposomes can be followed. The compositions of the present invention can also be formulated and used as tablets, capsules or elixirs for oral administration, suppositories for rectal administration, sterile solutions, suspensions for injectable administration, and the other compositions known in the art.

10 The present invention also includes pharmaceutically acceptable formulations of the compounds described. These formulations include salts of the above compounds, *e.g.*, acid addition salts, for example, salts of hydrochloric, hydrobromic, acetic acid, and benzene sulfonic acid.

15 A pharmacological composition or formulation refers to a composition or formulation in a form suitable for administration, *e.g.*, systemic administration, into a cell or subject, including for example a human. Suitable forms, in part, depend upon the use or the route of entry, for example oral, transdermal, or by injection. Such forms should not prevent the composition or formulation from reaching a target cell (*i.e.*, a cell to which the negatively charged nucleic acid is desirable for delivery). For example, pharmacological compositions injected into the blood stream should be soluble. Other factors are known in the art, and include considerations such as toxicity and forms that prevent the composition or formulation from exerting its effect.

20 By "systemic administration" is meant *in vivo* systemic absorption or accumulation of drugs in the blood stream followed by distribution throughout the entire body. Administration routes that lead to systemic absorption include, without limitation: intravenous, subcutaneous, intraperitoneal, inhalation, oral, intrapulmonary and intramuscular. Each of these administration routes exposes the siNA molecules of the invention to an accessible diseased tissue. The rate of entry of a drug into the circulation has been shown to be a function of molecular weight or size. The use of a liposome or other drug carrier comprising the compounds of the instant invention can potentially localize the drug, for example, in certain tissue types, such as the tissues of the reticular endothelial system (RES). A liposome formulation that can facilitate the association of drug with the surface of cells, such as, lymphocytes and macrophages is also useful. This approach can provide enhanced delivery of the drug to target cells by

30



taking advantage of the specificity of macrophage and lymphocyte immune recognition of abnormal cells, such as cells producing excess HIV.

By "pharmaceutically acceptable formulation" is meant, a composition or formulation that allows for the effective distribution of the nucleic acid molecules of the instant invention in the physical location most suitable for their desired activity. Non-limiting examples of agents suitable for formulation with the nucleic acid molecules of the instant invention include: P-glycoprotein inhibitors (such as Pluronic P85), which can enhance entry of drugs into the CNS (Joliet-Riant and Tillement, 1999, *Fundam. Clin. Pharmacol.*, 13, 16-26); biodegradable polymers, such as poly (DL-lactide-coglycolide) microspheres for sustained release delivery after intracerebral implantation (Emerich, DF *et al.*, 1999, *Cell Transplant*, 8, 47-58) (Alkermes, Inc. Cambridge, MA); and loaded nanoparticles, such as those made of polybutylcyanoacrylate, which can deliver drugs across the blood brain barrier and can alter neuronal uptake mechanisms (*Prog Neuropsychopharmacol Biol Psychiatry*, 23, 941-949, 1999). Other non-limiting examples of delivery strategies for the nucleic acid molecules of the instant invention include material described in Boado *et al.*, 1998, *J. Pharm. Sci.*, 87, 1308-1315; Tyler *et al.*, 1999, *FEBS Lett.*, 421, 280-284; Pardridge *et al.*, 1995, *PNAS USA.*, 92, 5592-5596; Boado, 1995, *Adv. Drug Delivery Rev.*, 15, 73-107; Aldrian-Herrada *et al.*, 1998, *Nucleic Acids Res.*, 26, 4910-4916; and Tyler *et al.*, 1999, *PNAS USA.*, 96, 7053-7058.

The invention also features the use of the composition comprising surface-modified liposomes containing poly (ethylene glycol) lipids (PEG-modified, or long-circulating liposomes or stealth liposomes). These formulations offer a method for increasing the accumulation of drugs in target tissues. This class of drug carriers resists opsonization and elimination by the mononuclear phagocytic system (MPS or RES), thereby enabling longer blood circulation times and enhanced tissue exposure for the encapsulated drug (Lasic *et al.* *Chem. Rev.* 1995, 95, 2601-2627; Ishiwata *et al.*, *Chem. Pharm. Bull.* 1995, 43, 1005-1011). Such liposomes have been shown to accumulate selectively in tumors, presumably by extravasation and capture in the neovascularized target tissues (Lasic *et al.*, *Science* 1995, 267, 1275-1276; Oku *et al.*, 1995, *Biochim. Biophys. Acta*, 1238, 86-90). The long-circulating liposomes enhance the pharmacokinetics and pharmacodynamics of DNA and RNA, particularly compared to conventional cationic liposomes which are known to accumulate in tissues of the MPS

(Liu *et al.*, *J. Biol. Chem.* 1995, 42, 24864-24870; Choi *et al.*, International PCT Publication No. WO 96/10391; Ansell *et al.*, International PCT Publication No. WO 96/10390; Holland *et al.*, International PCT Publication No. WO 96/10392). Long-circulating liposomes are also likely to protect drugs from nuclease degradation to a greater extent compared to cationic liposomes, based on their ability to avoid accumulation in metabolically aggressive MPS tissues such as the liver and spleen.

The present invention also includes compositions prepared for storage or administration that include a pharmaceutically effective amount of the desired compounds in a pharmaceutically acceptable carrier or diluent. Acceptable carriers or diluents for therapeutic use are well known in the pharmaceutical art, and are described, for example, in *Remington's Pharmaceutical Sciences*, Mack Publishing Co. (A.R. Gennaro edit. 1985), hereby incorporated by reference herein. For example, preservatives, stabilizers, dyes and flavoring agents can be provided. These include sodium benzoate, sorbic acid and esters of *p*-hydroxybenzoic acid. In addition, antioxidants and suspending agents can be used.

A pharmaceutically effective dose is that dose required to prevent, inhibit the occurrence, or treat (alleviate a symptom to some extent, preferably all of the symptoms) of a disease state. The pharmaceutically effective dose depends on the type of disease, the composition used, the route of administration, the type of mammal being treated, the physical characteristics of the specific mammal under consideration, concurrent medication, and other factors that those skilled in the medical arts will recognize. Generally, an amount between 0.1 mg/kg and 100 mg/kg body weight/day of active ingredients is administered dependent upon potency of the negatively charged polymer.

The nucleic acid molecules of the invention and formulations thereof can be administered orally, topically, parenterally, by inhalation or spray, or rectally in dosage unit formulations containing conventional non-toxic pharmaceutically acceptable carriers, adjuvants and/or vehicles. The term parenteral as used herein includes percutaneous, subcutaneous, intravascular (*e.g.*, intravenous), intramuscular, or intrathecal injection or infusion techniques and the like. In addition, there is provided a pharmaceutical formulation comprising a nucleic acid molecule of the invention and a pharmaceutically acceptable carrier. One or more nucleic acid molecules of the

invention can be present in association with one or more non-toxic pharmaceutically acceptable carriers and/or diluents and/or adjuvants, and if desired other active ingredients. The pharmaceutical compositions containing nucleic acid molecules of the invention can be in a form suitable for oral use, for example, as tablets, troches,  
5 lozenges, aqueous or oily suspensions, dispersible powders or granules, emulsion, hard or soft capsules, or syrups or elixirs.

Compositions intended for oral use can be prepared according to any method known to the art for the manufacture of pharmaceutical compositions and such compositions can contain one or more such sweetening agents, flavoring agents, coloring  
10 agents or preservative agents in order to provide pharmaceutically elegant and palatable preparations. Tablets contain the active ingredient in admixture with non-toxic pharmaceutically acceptable excipients that are suitable for the manufacture of tablets. These excipients can be, for example, inert diluents; such as calcium carbonate, sodium carbonate, lactose, calcium phosphate or sodium phosphate; granulating and  
15 disintegrating agents, for example, corn starch, or alginic acid; binding agents, for example starch, gelatin or acacia; and lubricating agents, for example magnesium stearate, stearic acid or talc. The tablets can be uncoated or they can be coated by known techniques. In some cases such coatings can be prepared by known techniques to delay disintegration and absorption in the gastrointestinal tract and thereby provide a sustained  
20 action over a longer period. For example, a time delay material such as glyceryl monostearate or glyceryl distearate can be employed.

Formulations for oral use can also be presented as hard gelatin capsules wherein the active ingredient is mixed with an inert solid diluent, for example, calcium carbonate, calcium phosphate or kaolin, or as soft gelatin capsules wherein the active ingredient is  
25 mixed with water or an oil medium, for example peanut oil, liquid paraffin or olive oil.

Aqueous suspensions contain the active materials in a mixture with excipients suitable for the manufacture of aqueous suspensions. Such excipients are suspending agents, for example sodium carboxymethylcellulose, methylcellulose, hydropropyl-methylcellulose, sodium alginate, polyvinylpyrrolidone, gum tragacanth and gum acacia;  
30 dispersing or wetting agents can be a naturally-occurring phosphatide, for example, lecithin, or condensation products of an alkylene oxide with fatty acids, for example

polyoxyethylene stearate, or condensation products of ethylene oxide with long chain aliphatic alcohols, for example heptadecaethyleneoxycetanol, or condensation products of ethylene oxide with partial esters derived from fatty acids and a hexitol such as polyoxyethylene sorbitol monooleate, or condensation products of ethylene oxide with partial esters derived from fatty acids and hexitol anhydrides, for example polyethylene sorbitan monooleate. The aqueous suspensions can also contain one or more preservatives, for example ethyl, or n-propyl p-hydroxybenzoate, one or more coloring agents, one or more flavoring agents, and one or more sweetening agents, such as sucrose or saccharin.

Oily suspensions can be formulated by suspending the active ingredients in a vegetable oil, for example arachis oil, olive oil, sesame oil or coconut oil, or in a mineral oil such as liquid paraffin. The oily suspensions can contain a thickening agent, for example beeswax, hard paraffin or cetyl alcohol. Sweetening agents and flavoring agents can be added to provide palatable oral preparations. These compositions can be preserved by the addition of an anti-oxidant such as ascorbic acid

Dispersible powders and granules suitable for preparation of an aqueous suspension by the addition of water provide the active ingredient in admixture with a dispersing or wetting agent, suspending agent and one or more preservatives. Suitable dispersing or wetting agents or suspending agents are exemplified by those already mentioned above. Additional excipients, for example sweetening, flavoring and coloring agents, can also be present.

Pharmaceutical compositions of the invention can also be in the form of oil-in-water emulsions. The oily phase can be a vegetable oil or a mineral oil or mixtures of these. Suitable emulsifying agents can be naturally-occurring gums, for example gum acacia or gum tragacanth, naturally-occurring phosphatides, for example soy bean, lecithin, and esters or partial esters derived from fatty acids and hexitol, anhydrides, for example sorbitan monooleate, and condensation products of the said partial esters with ethylene oxide, for example polyoxyethylene sorbitan monooleate. The emulsions can also contain sweetening and flavoring agents.

Syrups and elixirs can be formulated with sweetening agents, for example glycerol, propylene glycol, sorbitol, glucose or sucrose. Such formulations can also contain a

demulcent, a preservative and flavoring and coloring agents. The pharmaceutical compositions can be in the form of a sterile injectable aqueous or oleaginous suspension. This suspension can be formulated according to the known art using those suitable dispersing or wetting agents and suspending agents that have been mentioned above.

5 The sterile injectable preparation can also be a sterile injectable solution or suspension in a non-toxic parentally acceptable diluent or solvent, for example as a solution in 1,3-butanediol. Among the acceptable vehicles and solvents that can be employed are water, Ringer's solution and isotonic sodium chloride solution. In addition, sterile, fixed oils are conventionally employed as a solvent or suspending medium. For this purpose, any

10 bland fixed oil can be employed including synthetic mono-or diglycerides. In addition, fatty acids such as oleic acid find use in the preparation of injectables.

The nucleic acid molecules of the invention can also be administered in the form of suppositories, *e.g.*, for rectal administration of the drug. These compositions can be prepared by mixing the drug with a suitable non-irritating excipient that is solid at

15 ordinary temperatures but liquid at the rectal temperature and will therefore melt in the rectum to release the drug. Such materials include cocoa butter and polyethylene glycols.

Nucleic acid molecules of the invention can be administered parenterally in a sterile medium. The drug, depending on the vehicle and concentration used, can either

20 be suspended or dissolved in the vehicle. Advantageously, adjuvants such as local anesthetics, preservatives and buffering agents can be dissolved in the vehicle.

Dosage levels of the order of from about 0.1 mg to about 140 mg per kilogram of body weight per day are useful in the treatment of the above-indicated conditions (about 0.5 mg to about 7 g per subject per day). The amount of active ingredient that can be

25 combined with the carrier materials to produce a single dosage form varies depending upon the host treated and the particular mode of administration. Dosage unit forms generally contain between from about 1 mg to about 500 mg of an active ingredient.

It is understood that the specific dose level for any particular subject depends upon a variety of factors including the activity of the specific compound employed, the age,

30 body weight, general health, sex, diet, time of administration, route of administration,

and rate of excretion, drug combination and the severity of the particular disease undergoing therapy.

For administration to non-human animals, the composition can also be added to the animal feed or drinking water. It can be convenient to formulate the animal feed and  
5 drinking water compositions so that the animal takes in a therapeutically appropriate quantity of the composition along with its diet. It can also be convenient to present the composition as a premix for addition to the feed or drinking water.

The nucleic acid molecules of the present invention can also be administered to a subject in combination with other therapeutic compounds to increase the overall  
10 therapeutic effect. The use of multiple compounds to treat an indication can increase the beneficial effects while reducing the presence of side effects.

In one embodiment, the invention comprises compositions suitable for administering nucleic acid molecules of the invention to specific cell types. For example, the asialoglycoprotein receptor (ASGPr) (Wu and Wu, 1987, *J. Biol. Chem.*  
15 262, 4429-4432) is unique to hepatocytes and binds branched galactose-terminal glycoproteins, such as asialoorosomucoid (ASOR). In another example, the folate receptor is overexpressed in many cancer cells. Binding of such glycoproteins, synthetic glycoconjugates, or folates to the receptor takes place with an affinity that strongly depends on the degree of branching of the oligosaccharide chain, for example,  
20 triantennary structures are bound with greater affinity than biantennary or monoantennary chains (Baenziger and Fiete, 1980, *Cell*, 22, 611-620; Connolly *et al.*, 1982, *J. Biol. Chem.*, 257, 939-945). Lee and Lee, 1987, *Glycoconjugate J.*, 4, 317-328, obtained this high specificity through the use of N-acetyl-D-galactosamine as the carbohydrate moiety, which has higher affinity for the receptor, compared to galactose. This "clustering effect"  
25 has also been described for the binding and uptake of mannosyl-terminating glycoproteins or glycoconjugates (Ponpipom *et al.*, 1981, *J. Med. Chem.*, 24, 1388-1395). The use of galactose, galactosamine, or folate based conjugates to transport exogenous compounds across cell membranes can provide a targeted delivery approach to, for example, the treatment of liver disease, cancers of the liver, or other cancers. The  
30 use of bioconjugates can also provide a reduction in the required dose of therapeutic compounds required for treatment. Furthermore, therapeutic bioavailability,

pharmacodynamics, and pharmacokinetic parameters can be modulated through the use of nucleic acid bioconjugates of the invention. Non-limiting examples of such bioconjugates are described in Vargeese *et al.*, USSN 10/201,394, filed August 13, 2001; and Matulic-Adamic *et al.*, USSN 60/362,016, filed March 6, 2002.

5           Alternatively, certain siNA molecules of the instant invention can be expressed within cells from eukaryotic promoters (*e.g.*, Izant and Weintraub, 1985, *Science*, 229, 345; McGarry and Lindquist, 1986, *Proc. Natl. Acad. Sci.*, USA 83, 399; Scanlon *et al.*, 1991, *Proc. Natl. Acad. Sci. USA*, 88, 10591-5; Kashani-Sabet *et al.*, 1992, *Antisense Res. Dev.*, 2, 3-15; Dropulic *et al.*, 1992, *J. Virol.*, 66, 1432-41; Weerasinghe *et al.*,  
10   1991, *J. Virol.*, 65, 5531-4; Ojwang *et al.*, 1992, *Proc. Natl. Acad. Sci. USA*, 89, 10802-6; Chen *et al.*, 1992, *Nucleic Acids Res.*, 20, 4581-9; Sarver *et al.*, 1990 *Science*, 247, 1222-1225; Thompson *et al.*, 1995, *Nucleic Acids Res.*, 23, 2259; Good *et al.*, 1997, *Gene Therapy*, 4, 45. Those skilled in the art realize that any nucleic acid can be expressed in eukaryotic cells from the appropriate DNA/RNA vector. The activity of  
15   such nucleic acids can be augmented by their release from the primary transcript by a enzymatic nucleic acid (Draper *et al.*, PCT WO 93/23569, and Sullivan *et al.*, PCT WO 94/02595; Ohkawa *et al.*, 1992, *Nucleic Acids Symp. Ser.*, 27, 15-6; Taira *et al.*, 1991, *Nucleic Acids Res.*, 19, 5125-30; Ventura *et al.*, 1993, *Nucleic Acids Res.*, 21, 3249-55; Chowrira *et al.*, 1994, *J. Biol. Chem.*, 269, 25856.

20           In another aspect of the invention, RNA molecules of the present invention can be expressed from transcription units (see for example Couture *et al.*, 1996, *TIG*, 12, 510) inserted into DNA or RNA vectors. The recombinant vectors can be DNA plasmids or viral vectors. siNA expressing viral vectors can be constructed based on, but not limited to, adeno-associated virus, retrovirus, adenovirus, or alphavirus. In another embodiment,  
25   pol III based constructs are used to express nucleic acid molecules of the invention (see for example Thompson, U.S. Pats. Nos. 5,902,880 and 6,146,886). The recombinant vectors capable of expressing the siNA molecules can be delivered as described above, and persist in target cells. Alternatively, viral vectors can be used that provide for transient expression of nucleic acid molecules. Such vectors can be repeatedly  
30   administered as necessary. Once expressed, the siNA molecule interacts with the target mRNA and generates an RNAi response. Delivery of siNA molecule expressing vectors can be systemic, such as by intravenous or intra-muscular administration, by

administration to target cells ex-planted from a subject followed by reintroduction into the subject, or by any other means that would allow for introduction into the desired target cell (for a review see Couture *et al.*, 1996, *TIG.*, 12, 510).

In one aspect the invention features an expression vector comprising a nucleic acid  
5 sequence encoding at least one siNA molecule of the instant invention. The expression  
vector can encode one or both strands of a siNA duplex, or a single self-complementary  
strand that self hybridizes into a siNA duplex. The nucleic acid sequences encoding the  
siNA molecules of the instant invention can be operably linked in a manner that allows  
expression of the siNA molecule (see for example Paul *et al.*, 2002, *Nature*  
10 *Biotechnology*, 19, 505; Miyagishi and Taira, 2002, *Nature Biotechnology*, 19, 497; Lee  
*et al.*, 2002, *Nature Biotechnology*, 19, 500; and Novina *et al.*, 2002, *Nature Medicine*, 8,  
681-686.

In another aspect, the invention features an expression vector comprising: a) a  
transcription initiation region (*e.g.*, eukaryotic pol I, II or III initiation region); b) a  
15 transcription termination region (*e.g.*, eukaryotic pol I, II or III termination region); and  
c) a nucleic acid sequence encoding at least one of the siNA molecules of the instant  
invention, wherein said sequence is operably linked to said initiation region and said  
termination region in a manner that allows expression and/or delivery of the siNA  
molecule. The vector can optionally include an open reading frame (ORF) for a protein  
20 operably linked on the 5' side or the 3'-side of the sequence encoding the siNA of the  
invention and/or an intron (intervening sequences).

Transcription of the siNA molecule sequences can be driven from a promoter for  
eukaryotic RNA polymerase I (pol I), RNA polymerase II (pol II), or RNA polymerase  
III (pol III). Transcripts from pol II or pol III promoters are expressed at high levels in  
25 all cells; the levels of a given pol II promoter in a given cell type depends on the nature  
of the gene regulatory sequences (enhancers, silencers, etc.) present nearby. Prokaryotic  
RNA polymerase promoters are also used, providing that the prokaryotic RNA  
polymerase enzyme is expressed in the appropriate cells (Elroy-Stein and Moss, 1990,  
*Proc. Natl. Acad. Sci. U S A*, 87, 6743-7; Gao and Huang 1993, *Nucleic Acids Res.*, 21,  
30 2867-72; Lieber *et al.*, 1993, *Methods Enzymol.*, 217, 47-66; Zhou *et al.*, 1990, *Mol.*  
*Cell. Biol.*, 10, 4529-37). Several investigators have demonstrated that nucleic acid



molecules expressed from such promoters can function in mammalian cells (e.g. Kashani-Sabet *et al.*, 1992, *Antisense Res. Dev.*, 2, 3-15; Ojwang *et al.*, 1992, *Proc. Natl. Acad. Sci. U S A*, 89, 10802-6; Chen *et al.*, 1992, *Nucleic Acids Res.*, 20, 4581-9; Yu *et al.*, 1993, *Proc. Natl. Acad. Sci. U S A*, 90, 6340-4; L'Huillier *et al.*, 1992, *EMBO J.*, 11, 4411-8; Lisiewicz *et al.*, 1993, *Proc. Natl. Acad. Sci. U. S. A*, 90, 8000-4; Thompson *et al.*, 1995, *Nucleic Acids Res.*, 23, 2259; Sullenger & Cech, 1993, *Science*, 262, 1566). More specifically, transcription units such as the ones derived from genes encoding U6 small nuclear (snRNA), transfer RNA (tRNA) and adenovirus VA RNA are useful in generating high concentrations of desired RNA molecules such as siNA in cells (Thompson *et al.*, *supra*; Couture and Stinchcomb, 1996, *supra*; Noonberg *et al.*, 1994, *Nucleic Acid Res.*, 22, 2830; Noonberg *et al.*, U.S. Pat. No. 5,624,803; Good *et al.*, 1997, *Gene Ther.*, 4, 45; Beigelman *et al.*, International PCT Publication No. WO 96/18736. The above siNA transcription units can be incorporated into a variety of vectors for introduction into mammalian cells, including but not restricted to, plasmid DNA vectors, viral DNA vectors (such as adenovirus or adeno-associated virus vectors), or viral RNA vectors (such as retroviral or alphavirus vectors) (for a review see Couture and Stinchcomb, 1996, *supra*).

In another aspect the invention features an expression vector comprising a nucleic acid sequence encoding at least one of the siNA molecules of the invention in a manner that allows expression of that siNA molecule. The expression vector comprises in one embodiment: a) a transcription initiation region; b) a transcription termination region; and c) a nucleic acid sequence encoding at least one strand of the siNA molecule, wherein the sequence is operably linked to the initiation region and the termination region in a manner that allows expression and/or delivery of the siNA molecule.

In another embodiment the expression vector comprises: a) a transcription initiation region; b) a transcription termination region; c) an open reading frame; and d) a nucleic acid sequence encoding at least one strand of a siNA molecule, wherein the sequence is operably linked to the 3'-end of the open reading frame and wherein the sequence is operably linked to the initiation region, the open reading frame and the termination region in a manner that allows expression and/or delivery of the siNA molecule. In yet another embodiment, the expression vector comprises: a) a transcription initiation region; b) a transcription termination region; c) an intron; and d) a

nucleic acid sequence encoding at least one siNA molecule, wherein the sequence is operably linked to the initiation region, the intron and the termination region in a manner which allows expression and/or delivery of the nucleic acid molecule.

In another embodiment, the expression vector comprises: a) a transcription initiation region; b) a transcription termination region; c) an intron; d) an open reading frame; and e) a nucleic acid sequence encoding at least one strand of a siNA molecule, wherein the sequence is operably linked to the 3'-end of the open reading frame and wherein the sequence is operably linked to the initiation region, the intron, the open reading frame and the termination region in a manner which allows expression and/or delivery of the siNA molecule.

#### HIV overview

AIDS (acquired immunodeficiency syndrome) was first reported in the United States in 1981 and has since become a major worldwide epidemic. AIDS is caused by the human immunodeficiency virus (HIV). By killing or damaging cells of the body's immune system, HIV progressively destroys the body's ability to fight infections and certain cancers. People diagnosed with AIDS may get life-threatening diseases called opportunistic infections, which are caused by microbes such as viruses or bacteria that usually do not make healthy people sick. More than 790,000 cases of AIDS have been reported in the United States since 1981, and as many as 900,000 Americans may be infected with HIV.

HIV infection results in a chronic, progressive illness. Its course is marked by increasing levels of viral replication and the emergence of more virulent viral strains. This process causes the destruction of the immune system. HIV infection is staged by CD4 cell counts and clinical symptoms. Not all people progress through all stages and the time frames may also vary greatly from person to person. After infection with HIV, there is usually a seroconversion illness followed by an asymptomatic stage which lasts months to years (up to 18, but 10 years is the average). This is followed by symptomatic phases which correlate with progressive immunodeficiency. The progression of HIV disease varies from person to person and depends on a number of factors, including genetics and mode of transmission. Viral load is an important surrogate marker which measures the quantity of virus in plasma and predicts the rate of progression. It is a

measured in RNA copies/ml. It is also used to assess response to drug therapy and may predict the development of drug resistance. After seroconversion, each person develops a viral load set point. The lower the viral load the slower the progression of HIV disease (i.e. immune system destruction) and eventually clinical symptoms and opportunistic conditions. Between 50-90% of people infected with HIV have an acute clinical illness which typically occurs 2-4 weeks after the infecting exposure to HIV. Very often these seroconversion illnesses are recognized in retrospect since the symptoms may be non specific (viral or flu-like). Many develop a mononucleosis-like illness which begins acutely and lasts up to two weeks. Symptoms may include: fever, headache, lymphadenopathy, myalgia, rash, and mucocutaneous ulceration. Laboratory tests may indicate a viral infection. HIV p24 antigen may be detected at this stage even if antibodies to the HIV have not yet formed. Viral load is very high at this stage and CD4 counts drop transiently during seroconversion.

HIV is a retrovirus member of the Lentivirinae family. Retroviruses are enveloped RNA viruses characteristically possessing an RNA-dependent DNA polymerase termed reverse transcriptase. Two types of virus are known to affect humans; HIV-1 causes AIDS and is found worldwide, and HIV-2 has been isolated from some African cases of AIDS.

In its extracellular form, the virus exists as a lipid-enclosed cylindrical nucleocapsid of approximately 100 nm in diameter. Inserted within its lipid envelope are glycoproteins, a portion of which (glycoprotein [gp] 120) forms the binding region that attaches to the CD4 receptor on host cells (T-lymphocyte helper cells, activated monocytes and macrophages, and glial cells). After fusing with the cell membrane and entering the cytoplasm, the virus loses its envelope, and reverse transcription of RNA to DNA occurs.

Viral DNA integrates into host cell DNA as a latent provirus by a viral endonuclease. If the host cell is latently infected, no infection develops. If the host is actively infected, the proviral DNA is transcribed and translated, producing viral proteins and RNA. The viral proteins assemble and bud (by reverse endocytosis) through the host cell plasma membrane as new virions. The virus disseminates by budding or by cell-to-cell transfer.

- Abnormalities in all components of the immune system can be seen as the severity of HIV disease progresses. The most profound consequence of HIV infection is impairment of cell-mediated (T cell) immunity. HIV binds directly to the CD4 receptor of the T helper cell, resulting in progressive depletion of this T-cell population. As a result, the immune system is less able to (1) mount cytotoxic T-cell responses to virally infected cells or cancers, (2) to form delayed-type hypersensitivity reactions, and (3) to process new foreign substances presenting to the immune system. Significant impairment of the humoral immune system occurs in most persons infected with HIV. HIV also can infect monocytes and macrophages.
- 10 The use of small interfering nucleic acid molecules targeting HIV genes therefore provides a class of novel therapeutic agents that can be used in the treatment HIV infection and AIDS, as well as immunological disorders or any other disease or condition that responds to modulation of HIV genes.

Examples:

- 15 The following are non-limiting examples showing the selection, isolation, synthesis and activity of nucleic acids of the instant invention.

Example 1: Tandem synthesis of siNA constructs

- Exemplary siNA molecules of the invention are synthesized in tandem using a cleavable linker, for example, a succinyl-based linker. Tandem synthesis as described herein is followed by a one-step purification process that provides RNAi molecules in high yield. This approach is highly amenable to siNA synthesis in support of high throughput RNAi screening, and can be readily adapted to multi-column or multi-well synthesis platforms.
- 20

- After completing a tandem synthesis of a siNA oligo and its complement in which the 5'-terminal dimethoxytrityl (5'-O-DMT) group remains intact (trityl on synthesis), the oligonucleotides are deprotected as described above. Following deprotection, the siNA sequence strands are allowed to spontaneously hybridize. This hybridization yields a duplex in which one strand has retained the 5'-O-DMT group while the complementary strand comprises a terminal 5'-hydroxyl. The newly formed duplex behaves as a single molecule during routine solid-phase extraction purification (Trityl-On purification) even
- 25
- 30

though only one molecule has a dimethoxytrityl group. Because the strands form a stable duplex, this dimethoxytrityl group (or an equivalent group, such as other trityl groups or other hydrophobic moieties) is all that is required to purify the pair of oligos, for example by using a C18 cartridge.

5        Standard phosphoramidite synthesis chemistry is used up to the point of introducing a tandem linker, such as an inverted deoxy abasic succinate or glyceryl succinate linker (see **Figure 1**) or an equivalent cleavable linker. A non-limiting example of linker coupling conditions that can be used includes a hindered base such as diisopropylethylamine (DIPA) and/or DMAP in the presence of an activator reagent such  
10 as Bromotripyrrolidinophosphoniumhexafluorophosphate (PyBrOP). After the linker is coupled, standard synthesis chemistry is utilized to complete synthesis of the second sequence leaving the terminal the 5'-O-DMT intact. Following synthesis, the resulting oligonucleotide is deprotected according to the procedures described herein and quenched with a suitable buffer, for example with 50mM NaOAc or 1.5M  $\text{NH}_4\text{H}_2\text{CO}_3$ .

15        Purification of the siNA duplex can be readily accomplished using solid phase extraction, for example using a Waters C18 SepPak 1g cartridge conditioned with 1 column volume (CV) of acetonitrile, 2 CV  $\text{H}_2\text{O}$ , and 2 CV 50mM NaOAc. The sample is loaded and then washed with 1 CV  $\text{H}_2\text{O}$  or 50mM NaOAc. Failure sequences are eluted with 1 CV 14% ACN (Aqueous with 50mM NaOAc and 50mM NaCl). The  
20 column is then washed, for example with 1 CV  $\text{H}_2\text{O}$  followed by on-column detritylation, for example by passing 1 CV of 1% aqueous trifluoroacetic acid (TFA) over the column, then adding a second CV of 1% aqueous TFA to the column and allowing to stand for approximately 10 minutes. The remaining TFA solution is removed and the column washed with  $\text{H}_2\text{O}$  followed by 1 CV 1M NaCl and additional  
25  $\text{H}_2\text{O}$ . The siNA duplex product is then eluted, for example using 1 CV 20% aqueous CAN.

**Figure 2** provides an example of MALDI-TOV mass spectrometry analysis of a purified siNA construct in which each peak corresponds to the calculated mass of an individual siNA strand of the siNA duplex. The same purified siNA provides three  
30 peaks when analyzed by capillary gel electrophoresis (CGE), one peak presumably corresponding to the duplex siNA, and two peaks presumably corresponding to the

separate siNA sequence strands. Ion exchange HPLC analysis of the same siNA construct only shows a single peak. Testing of the purified siNA construct using a luciferase reporter assay described below demonstrated the same RNAi activity compared to siNA constructs generated from separately synthesized oligonucleotide sequence strands.

5    Example 2: Identification of potential siNA target sites in any RNA sequence

The sequence of an RNA target of interest, such as a viral or human mRNA transcript, is screened for target sites, for example by using a computer folding algorithm. In a non-limiting example, the sequence of a gene or RNA gene transcript derived from a database, such as Genbank, is used to generate siNA targets having  
10    complementarity to the target. Such sequences can be obtained from a database, or can be determined experimentally as known in the art. Target sites that are known, for example, those target sites determined to be effective target sites based on studies with other nucleic acid molecules, for example ribozymes or antisense, or those targets known to be associated with a disease or condition such as those sites containing mutations or  
15    deletions, can be used to design siNA molecules targeting those sites. Various parameters can be used to determine which sites are the most suitable target sites within the target RNA sequence. These parameters include but are not limited to secondary or tertiary RNA structure, the nucleotide base composition of the target sequence, the degree of homology between various regions of the target sequence, or the relative  
20    position of the target sequence within the RNA transcript. Based on these determinations, any number of target sites within the RNA transcript can be chosen to screen siNA molecules for efficacy, for example by using *in vitro* RNA cleavage assays, cell culture, or animal models. In a non-limiting example, anywhere from 1 to 1000 target sites are chosen within the transcript based on the size of the siNA construct to be  
25    used. High throughput screening assays can be developed for screening siNA molecules using methods known in the art, such as with multi-well or multi-plate assays to determine efficient reduction in target gene expression.

Example 3: Selection of siNA molecule target sites in a RNA

The following non-limiting steps can be used to carry out the selection of siNAs  
30    targeting a given gene sequence or transcript.

1. The target sequence is parsed *in silico* into a list of all fragments or subsequences of a particular length, for example 23 nucleotide fragments, contained within the target sequence. This step is typically carried out using a custom Perl script, but commercial sequence analysis programs such as Oligo, MacVector, or the GCG  
5 Wisconsin Package can be employed as well.
2. In some instances the siNAs correspond to more than one target sequence; such would be the case for example in targeting different transcripts of the same gene, targeting different transcripts of more than one gene, or for targeting both the human gene and an animal homolog. In this case, a subsequence list of a particular length is  
10 generated for each of the targets, and then the lists are compared to find matching sequences in each list. The subsequences are then ranked according to the number of target sequences that contain the given subsequence; the goal is to find subsequences that are present in most or all of the target sequences. Alternately, the ranking can identify subsequences that are unique to a target sequence, such as a mutant target  
15 sequence. Such an approach would enable the use of siNA to target specifically the mutant sequence and not effect the expression of the normal sequence.
3. In some instances the siNA subsequences are absent in one or more sequences while present in the desired target sequence; such would be the case if the siNA targets a gene with a paralogous family member that is to remain untargeted. As in case 2  
20 above, a subsequence list of a particular length is generated for each of the targets, and then the lists are compared to find sequences that are present in the target gene but are absent in the untargeted paralog.
4. The ranked siNA subsequences can be further analyzed and ranked according to GC content. A preference can be given to sites containing 30-70% GC, with a further  
25 preference to sites containing 40-60% GC.
5. The ranked siNA subsequences can be further analyzed and ranked according to self-folding and internal hairpins. Weaker internal folds are preferred; strong hairpin structures are to be avoided.
6. The ranked siNA subsequences can be further analyzed and ranked according to  
30 whether they have runs of GGG or CCC in the sequence. GGG (or even more Gs) in

either strand can make oligonucleotide synthesis problematic and can potentially interfere with RNAi activity, so it is avoided whenever better sequences are available. CCC is searched in the target strand because that will place GGG in the antisense strand.

- 5     7. The ranked siNA subsequences can be further analyzed and ranked according to whether they have the dinucleotide UU (uridine dinucleotide) on the 3'-end of the sequence, and/or AA on the 5'-end of the sequence (to yield 3' UU on the antisense sequence). These sequences allow one to design siNA molecules with terminal TT thymidine dinucleotides.
- 10    8. Four or five target sites are chosen from the ranked list of subsequences as described above. For example, in subsequences having 23 nucleotides, the right 21 nucleotides of each chosen 23-mer subsequence are then designed and synthesized for the upper (sense) strand of the siNA duplex, while the reverse complement of the left 21 nucleotides of each chosen 23-mer subsequence are then designed and synthesized  
 15    for the lower (antisense) strand of the siNA duplex (see **Tables II and III**). If terminal TT residues are desired for the sequence (as described in paragraph 7), then the two 3' terminal nucleotides of both the sense and antisense strands are replaced by TT prior to synthesizing the oligos.
- 20    9. The siNA molecules are screened in an *in vitro*, cell culture or animal model system to identify the most active siNA molecule or the most preferred target site within the target RNA sequence.

In an alternate approach, a pool of siNA constructs specific to a HIV target sequence is used to screen for target sites in cells expressing HIV RNA, such as in B cell, T cell, macrophage and endothelial cell culture systems. The general strategy used in  
 25    this approach is shown in **Figure 9**. A non-limiting example of such is a pool comprising sequences having sense sequences comprising SEQ ID NOs. 1-738, 1477-1490, 1499-1506, 1515-1522, 1531-1534, 1539-1546, 1555-1556, 1559-1566, and 1575-1576 and antisense sequences comprising SEQ ID NOs. 739-1476, 1491-1498, 1507-1514, 1523-1530, 1535-1538, 1547-1554, 1557-1558, and 1567-1574 respectively. Cells  
 30    expressing HIV (e.g., A549 or Caco-2 cells) are transfected with the pool of siNA constructs and cells that demonstrate a phenotype associated with HIV inhibition are



sorted. The pool of siNA constructs can be expressed from transcription cassettes inserted into appropriate vectors (see for example **Figure 7** and **Figure 8**). The siNA from cells demonstrating a positive phenotypic change (e.g., decreased proliferation, decreased HIV mRNA levels or decreased HIV protein expression), are sequenced to  
5 determine the most suitable target site(s) within the target HIV RNA sequence.

#### Example 4: HIV targeted siNA design

siNA target sites were chosen by analyzing sequences of the HIV RNA target and optionally prioritizing the target sites on the basis of folding (structure of any given sequence analyzed to determine siNA accessibility to the target), by using a library of  
10 siNA molecules as described in Example 3, or alternately by using an *in vitro* siNA system as described in Example 6 herein. siNA target sites were chosen by analyzing sequences of the HIV-1 RNA target (for example Genbank Accession Nos. shown in **Table III**) and optionally prioritizing the target sites on the basis of folding (structure of any given sequence analyzed to determine siNA accessibility to the target). The  
15 sequence alignments of all known A and B strains of HIV were screened for homology and siNA molecules were designed to target conserved sequences across these strains since the A and B strains are currently the most prevalent strains. Alternately, all known strains or other subclasses of HIV can be similarly screened for homology (see **Table IV**) and homologous sequences used as targets. A cutoff for % homology between the  
20 different strains can be used to increase or decrease the number of targets considered, for example 70%, 75%, 80%, 85%, 90% or 95% homology. The sequences shown in **Table I** represent 80% homology between the HIV strains shown in **Table III**. siNA molecules were designed that could bind each target sequence and are optionally individually analyzed by computer folding to assess whether the siNA molecule can interact with the  
25 target sequence. Varying the length of the siNA molecules can be chosen to optimize activity. The siNA sense (upper sequence) and antisense (lower sequence) sequences shown in **Table I** comprise 19 nucleotides in length, with the sense strand comprising the same sequence as the target sequence and the antisense strand comprising a complementary sequence to the sense/target sequence. The sense and antisense strands  
30 can further comprise nucleotide 3'-overhangs as described herein, preferably the overhangs comprise about 2 nucleotides which can optionally be complementary to the target sequence in the antisense siNA strand, and/or optionally analogous to the adjacent

nucleotides in the target sequence when present in the sense siNA strand. Generally, a sufficient number of complementary nucleotide bases are chosen to bind to, or otherwise interact with, the target RNA, but the degree of complementarity can be modulated to accommodate siNA duplexes of varying length or base composition. By using such methodologies, siNA molecules can be designed to target sites within any known RNA sequence, for example those RNA sequences corresponding to the any gene transcript.

Chemically modified siNA constructs are designed to provide nuclease stability for systemic administration in vivo and/or improved pharmacokinetic, localization, and delivery properties while preserving the ability to mediate RNAi activity. Chemical modifications as described herein are introduced synthetically using synthetic methods described herein and those generally known in the art. The synthetic siNA constructs are then assayed for nuclease stability in serum and/or cellular/tissue extracts (e.g. liver extracts). The synthetic siNA constructs are also tested in parallel for RNAi activity using an appropriate assay, such as a luciferase reporter assay as described herein or another suitable assay that can quantify RNAi activity. Synthetic siNA constructs that possess both nuclease stability and RNAi activity can be further modified and re-evaluated in stability and activity assays. The chemical modifications of the stabilized active siNA constructs can then be applied to any siNA sequence targeting any chosen RNA and used, for example, in target screening assays to pick lead siNA compounds for therapeutic development (see for example **Figure 11**).

#### Example 5: Chemical Synthesis and Purification of siNA

siNA molecules can be designed to interact with various sites in the RNA message, for example, target sequences within the RNA sequences described herein. The sequence of one strand of the siNA molecule(s) is complementary to the target site sequences described above. The siNA molecules can be chemically synthesized using methods described herein. Inactive siNA molecules that are used as control sequences can be synthesized by scrambling the sequence of the siNA molecules such that it is not complementary to the target sequence. Generally, siNA constructs can be synthesized using solid phase oligonucleotide synthesis methods as described herein (see for example Usman *et al.*, US Patent Nos. 5,804,683; 5,831,071; 5,998,203; 6,117,657; 6,353,098; 6,362,323; 6,437,117; 6,469,158; Scaringe *et al.*, US Patent Nos. 6,111,086; 6,008,400; 6,111,086 all incorporated by reference herein in their entirety).

In a non-limiting example, RNA oligonucleotides are synthesized in a stepwise fashion using the phosphoramidite chemistry as is known in the art. Standard phosphoramidite chemistry involves the use of nucleosides comprising any of 5'-O-dimethoxytrityl, 2'-O-tert-butyldimethylsilyl, 3'-O-2-Cyanoethyl N,N-diisopropylphosphoroamidite groups, and exocyclic amine protecting groups (e.g. N6-benzoyl adenosine, N4 acetyl cytidine, and N2-isobutyryl guanosine). Alternately, 2'-O-Silyl Ethers can be used in conjunction with acid-labile 2'-O-orthoester protecting groups in the synthesis of RNA as described by Scaringe *supra*. Differing 2' chemistries can require different protecting groups, for example 2'-deoxy-2'-amino nucleosides can utilize N-phthaloyl protection as described by Usman *et al.*, US Patent 5,631,360, incorporated by reference herein in its entirety).

During solid phase synthesis, each nucleotide is added sequentially (3'- to 5'-direction) to the solid support-bound oligonucleotide. The first nucleoside at the 3'-end of the chain is covalently attached to a solid support (e.g., controlled pore glass or polystyrene) using various linkers. The nucleotide precursor, a ribonucleoside phosphoramidite, and activator are combined resulting in the coupling of the second nucleoside phosphoramidite onto the 5'-end of the first nucleoside. The support is then washed and any unreacted 5'-hydroxyl groups are capped with a capping reagent such as acetic anhydride to yield inactive 5'-acetyl moieties. The trivalent phosphorus linkage is then oxidized to a more stable phosphate linkage. At the end of the nucleotide addition cycle, the 5'-O-protecting group is cleaved under suitable conditions (e.g., acidic conditions for trityl-based groups and Fluoride for silyl-based groups). The cycle is repeated for each subsequent nucleotide.

Modification of synthesis conditions can be used to optimize coupling efficiency, for example by using differing coupling times, differing reagent/phosphoramidite concentrations, differing contact times, differing solid supports and solid support linker chemistries depending on the particular chemical composition of the siNA to be synthesized. Deprotection and purification of the siNA can be performed as is generally described in Usman *et al.*, US 5,831,071, US 6,353,098, US 6,437,117, and Bellon *et al.*, US 6,054,576, US 6,162,909, US 6,303,773, incorporated by reference herein in their entirety or Scaringe *supra*. Additionally, deprotection conditions can be modified to provide the best possible yield and purity of siNA constructs. For example, applicant has

observed that oligonucleotides comprising 2'-deoxy-2'-fluoro nucleotides can degrade under inappropriate deprotection conditions. Such oligonucleotides are deprotected using aqueous methylamine at about 35°C for 30 minutes. If the 2'-deoxy-2'-fluoro containing oligonucleotide also comprises ribonucleotides, after deprotection with aqueous methylamine at about 35°C for 30 minutes, TEA-HF is added and the reaction maintained at about 65°C for an additional 15 minutes.

Example 6: RNAi *in vitro* assay to assess siNA activity

An *in vitro* assay that recapitulates RNAi in a cell-free system is used to evaluate siNA constructs targeting HIV RNA targets. The assay comprises the system described by Tuschl *et al.*, 1999, *Genes and Development*, 13, 3191-3197 and Zamore *et al.*, 2000, *Cell*, 101, 25-33 adapted for use with HIV target RNA. A *Drosophila* extract derived from syncytial blastoderm is used to reconstitute RNAi activity *in vitro*. Target RNA is generated via *in vitro* transcription from an appropriate HIV expressing plasmid using T7 RNA polymerase or via chemical synthesis as described herein. Sense and antisense siNA strands (for example 20 uM each) are annealed by incubation in buffer (such as 100 mM potassium acetate, 30 mM HEPES-KOH, pH 7.4, 2 mM magnesium acetate) for 1 min. at 90°C followed by 1 hour at 37°C, then diluted in lysis buffer (for example 100 mM potassium acetate, 30 mM HEPES-KOH at pH 7.4, 2mM magnesium acetate). Annealing can be monitored by gel electrophoresis on an agarose gel in TBE buffer and stained with ethidium bromide. The *Drosophila* lysate is prepared using zero to two-hour-old embryos from Oregon R flies collected on yeasted molasses agar that are dechorionated and lysed. The lysate is centrifuged and the supernatant isolated. The assay comprises a reaction mixture containing 50% lysate [vol/vol], RNA (10-50 pM final concentration), and 10% [vol/vol] lysis buffer containing siNA (10 nM final concentration). The reaction mixture also contains 10 mM creatine phosphate, 10 ug/ml creatine phosphokinase, 100 uM GTP, 100 uM UTP, 100 uM CTP, 500 uM ATP, 5 mM DTT, 0.1 U/uL RNasin (Promega), and 100 uM of each amino acid. The final concentration of potassium acetate is adjusted to 100 mM. The reactions are pre-assembled on ice and preincubated at 25° C for 10 minutes before adding RNA, then incubated at 25° C for an additional 60 minutes. Reactions are quenched with 4 volumes of 1.25 x Passive Lysis Buffer (Promega). Target RNA cleavage is assayed by RT-PCR

analysis or other methods known in the art and are compared to control reactions in which siNA is omitted from the reaction.

Alternately, internally-labeled target RNA for the assay is prepared by *in vitro* transcription in the presence of [ $\alpha$ - $^{32}$ P] CTP, passed over a G 50 Sephadex column  
5 by spin chromatography and used as target RNA without further purification. Optionally, target RNA is 5'- $^{32}$ P-end labeled using T4 polynucleotide kinase enzyme. Assays are performed as described above and target RNA and the specific RNA cleavage products generated by RNAi are visualized on an autoradiograph of a gel. The percentage of cleavage is determined by Phosphor Imager<sup>®</sup> quantitation of bands  
10 representing intact control RNA or RNA from control reactions without siNA and the cleavage products generated by the assay.

In one embodiment, this assay is used to determine target sites the HIV RNA target for siNA mediated RNAi cleavage, wherein a plurality of siNA constructs are screened for RNAi mediated cleavage of the HIV RNA target, for example, by analyzing the assay  
15 reaction by electrophoresis of labeled target RNA, or by northern blotting, as well as by other methodology well known in the art.

#### Example 7: Nucleic acid inhibition of HIV target RNA *in vivo*

siNA molecules targeted to the huma HIV RNA are designed and synthesized as described above. These nucleic acid molecules can be tested for cleavage activity *in*  
20 *vivo*, for example, using the following procedure. The target sequences and the nucleotide location within the HIV RNA are given in Table II and III.

Two formats are used to test the efficacy of siNAs targeting HIV. First, the reagents are tested in cell culture using, for example, B cell, T cell, macrophage or endothelial cell culture systems to determine the extent of RNA and protein inhibition.  
25 siNA reagents (*e.g.*; see Tables II and III) are selected against the HIV target as described herein. RNA inhibition is measured after delivery of these reagents by a suitable transfection agent to, for example, B cell, T cell, macrophage or endothelial cells. Relative amounts of target RNA are measured versus actin using real-time PCR monitoring of amplification (*eg.*, ABI 7700 Taqman<sup>®</sup>). A comparison is made to a  
30 mixture of oligonucleotide sequences made to unrelated targets or to a randomized siNA

control with the same overall length and chemistry, but randomly substituted at each position. Primary and secondary lead reagents are chosen for the target and optimization performed. After an optimal transfection agent concentration is chosen, a RNA time-course of inhibition is performed with the lead siNA molecule. In addition, a cell-plating  
5 format can be used to determine RNA inhibition.

#### Delivery of siNA to Cells

Cells (e.g., B cell, T cell, macrophage or endothelial cells) are seeded, for example, at  $1 \times 10^5$  cells per well of a six-well dish in EGM-2 (BioWhittaker) the day before transfection. siNA (final concentration, for example 20nM) and cationic lipid (e.g., final  
10 concentration  $2 \mu\text{g/ml}$ ) are complexed in EGM basal media (Biowhittaker) at  $37^\circ\text{C}$  for 30 minutes in polystyrene tubes. Following vortexing, the complexed siNA is added to each well and incubated for the times indicated. For initial optimization experiments, cells are seeded, for example, at  $1 \times 10^3$  in 96 well plates and siNA complex added as described. Efficiency of delivery of siNA to cells is determined using a fluorescent siNA complexed  
15 with lipid. Cells in 6-well dishes are incubated with siNA for 24 hours, rinsed with PBS and fixed in 2% paraformaldehyde for 15 minutes at room temperature. Uptake of siNA is visualized using a fluorescent microscope.

#### Taqman and Lightcycler quantification of mRNA

Total RNA is prepared from cells following siNA delivery, for example, using  
20 Qiagen RNA purification kits for 6-well or Rneasy extraction kits for 96-well assays. For Taqman analysis, dual-labeled probes are synthesized with the reporter dye, FAM or JOE, covalently linked at the 5'-end and the quencher dye TAMRA conjugated to the 3'-end. One-step RT-PCR amplifications are performed on, for example, an ABI PRISM 7700 Sequence Detector using 50  $\mu\text{l}$  reactions consisting of 10  $\mu\text{l}$  total RNA, 100 nM  
25 forward primer, 900 nM reverse primer, 100 nM probe, 1X TaqMan PCR reaction buffer (PE-Applied Biosystems), 5.5 mM  $\text{MgCl}_2$ , 300  $\mu\text{M}$  each dATP, dCTP, dGTP, and dTTP, 10U RNase Inhibitor (Promega), 1.25U AmpliTaq Gold (PE-Applied Biosystems) and 10U M-MLV Reverse Transcriptase (Promega). The thermal cycling conditions can consist of 30 min at  $48^\circ\text{C}$ , 10 min at  $95^\circ\text{C}$ , followed by 40 cycles of 15 sec at  $95^\circ\text{C}$  and 1  
30 min at  $60^\circ\text{C}$ . Quantitation of mRNA levels is determined relative to standards generated from serially diluted total cellular RNA (300, 100, 33, 11 ng/rxn) and normalizing to  $\beta$ -

actin or GAPDH mRNA in parallel TaqMan reactions. For each gene of interest an upper and lower primer and a fluorescently labeled probe are designed. Real time incorporation of SYBR Green I dye into a specific PCR product can be measured in glass capillary tubes using a lightcycler. A standard curve is generated for each primer pair  
5 using control cRNA. Values are represented as relative expression to GAPDH in each sample.

#### Western blotting

Nuclear extracts can be prepared using a standard micro preparation technique (see for example Andrews and Faller, 1991, *Nucleic Acids Research*, 19, 2499). Protein  
10 extracts from supernatants are prepared, for example using TCA precipitation. An equal volume of 20% TCA is added to the cell supernatant, incubated on ice for 1 hour and pelleted by centrifugation for 5 minutes. Pellets are washed in acetone, dried and resuspended in water. Cellular protein extracts are run on a 10% Bis-Tris NuPage (nuclear extracts) or 4-12% Tris-Glycine (supernatant extracts) polyacrylamide gel and  
15 transferred onto nitro-cellulose membranes. Non-specific binding can be blocked by incubation, for example, with 5% non-fat milk for 1 hour followed by primary antibody for 16 hour at 4°C. Following washes, the secondary antibody is applied, for example (1:10,000 dilution) for 1 hour at room temperature and the signal detected with SuperSignal reagent (Pierce).

#### 20 Example 8: Models useful to evaluate the down-regulation of HIV gene expression

##### *Cell Culture*

The siNA constructs of the invention can be used in various cell culture systems as are commonly known in the art to screen for compounds having anti-HIV activity. B cell, T cell, macrophage and endothelial cell culture systems are non-limiting examples  
25 of cell culture systems that can be readily adapted for screening siNA molecules of the invention. In a non-limiting example, siNA molecules of the invention are co-transfected with HIV-1 pNL4-3 proviral DNA into 293/EcR cells as described by Lee *et al.*, 2002, *Nature Biotechnology*, 19, 500-505, using a U6 snRNA promoter driven expression system.

In a non-limiting example, the siNA expression vectors are prepared using the pTZ U6+1 vector described in Lee *et al. supra.* as follows. One cassette harbors the 21-nucleotide sense sequences and the other a 21-nucleotide antisense sequence (**Table I**). These sequences are designed to target HIV-1 RNA targets described herein. As a control to verify a siNA mechanism, irrelevant sense and antisense (S/AS) sequences lacking complementarity to HIV-1 (S/AS (IR)) are subcloned in pTZ U6+1. RNA samples are prepared from 293/EcR cells transiently co-transfected with siNA or control constructs, and subjected to Ponasterone A induction. RNAs are also prepared from 293 cells co-transfected with HIV-1 pNL4-3 proviral DNA and siNA or control constructs. For determination of anti-HIV-1 activity of the siNAs, transient assays are done by co-transfection of siNA constructs and infectious HIV-1 proviral DNA, pNL4-3 into 293 cells as described above, followed by Northern analysis as known in the art. The p24 values are calculated with the aid of, for example, a Dynatech MR5000 ELISA plate reader (Dynatech Labs Inc., Chantilly, VA). Cell viability can also be assessed using a Trypan Blue dye exclusion count at four days after transfection.

Other cell culture model systems for screening compounds having anti-HIV activity are generally known in the art. For example, Duzgunes *et al.*, 2001, *Nucleosides, Nucleotides & Nucleic Acids*, 20(4-7), 515-523; Cagnun *et al.*, 2000, *Antisense Nucleic Acid Drug Dev.*, 10, 251; Ho *et al.*, 1995, *Stem Cells*, 13 supp 3, 100; and Baur *et al.*, 1997, *Blood*, 89, 2259 describe cell culture systems that can be readily adapted for use with the compositions of the instant invention and the assays described herein.

#### *Animal Models*

Evaluating the efficacy of anti-HIV agents in animal models is an important prerequisite to human clinical trials. The siNA constructs of the invention can be evaluated in a variety of animal models including, for example, a hollow fiber HIV model (see, for example, Gruenberg, US 5,627,070), mouse models for AIDS using transgenic mice expressing HIV-1 genes from CD4 promoters and enhancers (see, for example, Jolicoeur, International PCT Publication No. WO 98/50535) and/or the HIV/SIV/SHIV non-human primate models (see, for example, Narayan, US 5,849,994). The siNA compounds and virus can be administered by a variety of methods and routes as described herein and as known in the art. Quantitation of results in these models can



be performed by a variety of methods, including quantitative PCR, quantitative and bulk co-cultivation assays, plasma co-cultivation assays, antigen and antibody detection assays, lymphocyte proliferation, intracellular cytokines, flow cytometry, as well as hematology and CBC evaluation. Additional animal models are generally known in the art, see for example Bai *et al.*, 2000, *Mol. Ther.*, 1, 244.

Example 9: RNAi mediated inhibition of HIV RNA expression

siNA constructs (Table III) are tested for efficacy in reducing HIV RNA expression in, for example, 293 cells. Cells are plated approximately 24h before transfection in 96-well plates at 5,000-7,500 cells/well, 100  $\mu$ l/well, such that at the time of transfection cells are 70-90% confluent. For transfection, annealed siNAs are mixed with the transfection reagent in a volume of 50  $\mu$ l/well and incubated for 20 minutes at room temperature and are cotransfected with HIV-1 pNL4-3 proviral DNA into 293 cells. The siNA transfection mixtures are added to cells to give a final siNA concentration of 25 nM in a volume of 150  $\mu$ l. Each siNA transfection mixture is added to 3 wells for triplicate siNA treatments. Cells are incubated at 37° for 24h in the continued presence of the siNA transfection mixture. At 24h, RNA is prepared from each well of treated cells. The supernatants with the transfection mixtures are first removed and discarded, then the cells are lysed and RNA prepared from each well. Target gene expression following treatment is evaluated by RT-PCR for the target gene and for a control gene (36B4, an RNA polymerase subunit) for normalization. In addition, ELISA can be used to determine HIV-1 p24 viral antigen levels. The triplicate data is averaged and the standard deviations determined for each treatment. Normalized data are graphed and the percent reduction of target mRNA by active siNAs in comparison to their respective inverted control siNAs is determined (see for example Lee *et al.*, 2002, *Nature Biotechnology*, 20, 500-505).

In a non-limiting example, a siNA construct comprising ribonucleotides and 3'-terminal dithymidine caps is assayed along with a chemically modified siNA construct comprising 2'-deoxy-2'-fluoro pyrimidine nucleotides and purine ribonucleotides in which the sense strand of the siNA is further modified with 5' and 3'-terminal inverted deoxyabasic caps and the antisense strand comprises a 3'-terminal phosphorothioate internucleotide linkage. Additional stabilization chemistries as described in Table IV are

similarly assayed for activity. These siNA constructs are compared to appropriate matched chemistry inverted controls. In addition, the siNA constructs are also compared to untreated cells, cells transfected with lipid and scrambled siNA constructs, and cells transfected with lipid alone (transfection control).

5    Example 10: Indications

The present body of knowledge in HIV research indicates the need for methods to assay HIV activity and for compounds that can regulate HIV expression for research, diagnostic, and therapeutic use. As described herein, the nucleic acid molecules of the present invention can be used in assays to diagnose disease state related of HIV levels.

10   In addition, the nucleic acid molecules can be used to treat disease state related to HIV levels.

Particular degenerative and disease states that can be associated with HIV expression modulation include, but are not limited to, acquired immunodeficiency disease (AIDS) and related diseases and conditions including, but not limited to,

15   Kaposi's sarcoma, lymphoma, cervical cancer, squamous cell carcinoma, cardiac myopathy, rheumatic diseases, and opportunistic infection, for example Pneumocystis carinii, Cytomegalovirus, Herpes simplex, Mycobacteria, Cryptococcus, Toxoplasma, Progressive multifocal leuco-encephalopathy (Papovavirus), Mycobacteria, Aspergillus, Cryptococcus, Candida, Cryptosporidium, Isospora belli, Microsporidia and any other

20   diseases or conditions that are related to or will respond to the levels of HIV in a cell or tissue, alone or in combination with other therapies.

The present body of knowledge in HIV research indicates the need for methods to assay HIV activity and for compounds that can regulate HIV expression for research, diagnostic, and therapeutic use.

25       The use of antiviral compounds, monoclonal antibodies, chemotherapy, radiation therapy, analgesics, and/or anti-inflammatory compounds, are all non-limiting examples of a methods that can be combined with or used in conjunction with the nucleic acid molecules (e.g. ribozymes and antisense molecules) of the instant invention. Examples of antiviral compounds that can be used in conjunction with the nucleic acid molecules of

30   the invention include, but are not limited to, AZT (also known as zidovudine or ZDV),

ddC (zalcitabine), ddI (dideoxyinosine), d4T (stavudine), and 3TC (lamivudine) Ribavirin, delvaridine (Rescriptor), nevirapine (Viramune), efavirenz (Sustiva), ritonavir (Norvir), saquinavir (Invirase), indinavir (Crixivan), amprenavir (Agenerase), nelfinavir (Viracept), and/or lopinavir (Kaletra). Common chemotherapies that can be  
5 combined with nucleic acid molecules of the instant invention include various combinations of cytotoxic drugs to kill cancer cells. These drugs include, but are not limited to, paclitaxel (Taxol), docetaxel, cisplatin, methotrexate, cyclophosphamide, doxorubin, fluorouracil carboplatin, edatrexate, gemcitabine, and vinorelbine. Those skilled in the art will recognize that other drug compounds and therapies can be similarly  
10 be readily combined with the nucleic acid molecules of the instant invention (e.g. ribozymes, siRNA and antisense molecules) are hence within the scope of the instant invention.

#### Example 11: Diagnostic uses

The siNA molecules of the invention can be used in a variety of diagnostic  
15 applications, such as in the identification of molecular targets (e.g., RNA) in a variety of applications, for example, in clinical, industrial, environmental, agricultural and/or research settings. Such diagnostic use of siNA molecules involves utilizing reconstituted RNAi systems, for example, using cellular lysates or partially purified cellular lysates. siNA molecules of this invention can be used as diagnostic tools to examine genetic drift  
20 and mutations within diseased cells or to detect the presence of endogenous or exogenous, for example viral, RNA in a cell. The close relationship between siNA activity and the structure of the target RNA allows the detection of mutations in any region of the molecule, which alters the base-pairing and three-dimensional structure of the target RNA. By using multiple siNA molecules described in this invention, one can  
25 map nucleotide changes, which are important to RNA structure and function *in vitro*, as well as in cells and tissues. Cleavage of target RNAs with siNA molecules can be used to inhibit gene expression and define the role of specified gene products in the progression of disease or infection. In this manner, other genetic targets can be defined as important mediators of the disease. These experiments will lead to better treatment of  
30 the disease progression by affording the possibility of combination therapies (e.g., multiple siNA molecules targeted to different genes, siNA molecules coupled with known small molecule inhibitors, or intermittent treatment with combinations siNA

molecules and/or other chemical or biological molecules). Other *in vitro* uses of siNA molecules of this invention are well known in the art, and include detection of the presence of mRNAs associated with a disease, infection, or related condition. Such RNA is detected by determining the presence of a cleavage product after treatment with a  
5 siNA using standard methodologies, for example, fluorescence resonance emission transfer (FRET).

In a specific example, siNA molecules that cleave only wild-type or mutant forms of the target RNA are used for the assay. The first siNA molecules (*i.e.*, those that cleave only wild-type forms of target RNA) are used to identify wild-type RNA present  
10 in the sample and the second siNA molecules (*i.e.*, those that cleave only mutant forms of target RNA) are used to identify mutant RNA in the sample. As reaction controls, synthetic substrates of both wild-type and mutant RNA are cleaved by both siNA molecules to demonstrate the relative siNA efficiencies in the reactions and the absence of cleavage of the "non-targeted" RNA species. The cleavage products from the  
15 synthetic substrates also serve to generate size markers for the analysis of wild-type and mutant RNAs in the sample population. Thus, each analysis requires two siNA molecules, two substrates and one unknown sample, which is combined into six reactions. The presence of cleavage products is determined using an RNase protection assay so that full-length and cleavage fragments of each RNA can be analyzed in one  
20 lane of a polyacrylamide gel. It is not absolutely required to quantify the results to gain insight into the expression of mutant RNAs and putative risk of the desired phenotypic changes in target cells. The expression of mRNA whose protein product is implicated in the development of the phenotype (*i.e.*, disease related or infection related) is adequate to establish risk. If probes of comparable specific activity are used for both transcripts,  
25 then a qualitative comparison of RNA levels is adequate and decreases the cost of the initial diagnosis. Higher mutant form to wild-type ratios are correlated with higher risk whether RNA levels are compared qualitatively or quantitatively.

All patents and publications mentioned in the specification are indicative of the levels of skill of those skilled in the art to which the invention pertains. All references  
30 cited in this disclosure are incorporated by reference to the same extent as if each reference had been incorporated by reference in its entirety individually.

One skilled in the art would readily appreciate that the present invention is well adapted to carry out the objects and obtain the ends and advantages mentioned, as well as those inherent therein. The methods and compositions described herein as presently representative of preferred embodiments are exemplary and are not intended as  
5 limitations on the scope of the invention. Changes therein and other uses will occur to those skilled in the art, which are encompassed within the spirit of the invention, are defined by the scope of the claims.

It will be readily apparent to one skilled in the art that varying substitutions and modifications can be made to the invention disclosed herein without departing from the  
10 scope and spirit of the invention. Thus, such additional embodiments are within the scope of the present invention and the following claims. The present invention teaches one skilled in the art to test various combinations and/or substitutions of chemical modifications described herein toward generating nucleic acid constructs with improved activity for mediating RNAi activity. Such improved activity can comprise improved  
15 stability, improved bioavailability, and/or improved activation of cellular responses mediating RNAi. Therefore, the specific embodiments described herein are not limiting and one skilled in the art can readily appreciate that specific combinations of the modifications described herein can be tested without undue experimentation toward identifying siNA molecules with improved RNAi activity.

20 The invention illustratively described herein suitably can be practiced in the absence of any element or elements, limitation or limitations that are not specifically disclosed herein. Thus, for example, in each instance herein any of the terms "comprising", "consisting essentially of", and "consisting of" may be replaced with either of the other two terms. The terms and expressions which have been employed are used  
25 as terms of description and not of limitation, and there is no intention that in the use of such terms and expressions of excluding any equivalents of the features shown and described or portions thereof, but it is recognized that various modifications are possible within the scope of the invention claimed. Thus, it should be understood that although the present invention has been specifically disclosed by preferred embodiments, optional  
30 features, modification and variation of the concepts herein disclosed may be resorted to by those skilled in the art, and that such modifications and variations are considered to be within the scope of this invention as defined by the description and the appended claims.

In addition, where features or aspects of the invention are described in terms of Markush groups or other grouping of alternatives, those skilled in the art will recognize that the invention is also thereby described in terms of any individual member or subgroup of members of the Markush group or other group.

**Table I: HIV Accession Numbers**

<b>Accession</b>	<b>Name</b>	<b>Subtype</b>
AB032740	95TNIH022	01_AE
AB032741	95TNIH047	01_AE
AB052995	93JPNH1	01_AE
AB070352	NH25 93JPNH25T 93JP-NH2.5T	01_AE
AB070353	NH2 93JPNH2ENV	01_AE
AF164485	93TH9021	01_AE
AF197338	93TH057	01_AE
AF197339	93TH065	01_AE
AF197340	90CF11697 AF197340	01_AE
AF197341	90CF4071 AF197341	01_AE
AF259954	CM235-2	01_AE
AF259955	CM235-4	01_AE
AY008714	97CNGX2F 97CNGX-2F	01_AE
AY008718	97CNGX11F	01_AE
U51188	90CF402 90CR402 CAR-E 4002	01_AE
U51189	93TH253	01_AE
U54771	CM240	01_AE
AF362994	NP1623	01B
AY082968	TH1326 AY082968	01B
AJ404325	97DCKTB49 97CDKTB49 HIM404325	01GHJKU
AB049811	97GHAG1 AB049811	02_AG
AB052867	AB052867	02_AG
AF063223	DJ263	02_AG
AF063224	DJ264	02_AG
AF107770	SE7812	02_AG
AF184155	G829	02_AG
AF377954	CM52885 AF377954	02_AG
AF377955	CM53658 AF377955	02_AG
AJ251056	MP1211 98SE-MP1211	02_AG
AJ251057	MP1213 98SEMP1213 HIM251057	02_AG
AJ286133	97CM-MP807	02_AG
L39106	IBNG	02_AG
AF193276	KAL153-2	03_AB
AF193277	RU98001 98RU001	03_AB
AF414006	98BY10443 AF414006	03-AB
AF049337	94CY032-3 CY032.3	04_cpx
AF119819	97PVMY GR84	04_cpx
AF119820	97PVCH GR11	04_cpx
AF076998	VI961	05_DF
AF193253	VI1310 AF193253	05_DF
AF064699	BFP90	06_cpx
AJ245481	95ML84	06_cpx
AJ288981	97SE1078	06_cpx
AJ288982	95ML127	06_cpx
AF286226	97CN001 C54	07_BC
AF286230	98CN009	07_BC
AX149647	C54A C54	07_BC
AX149672	C54D AX149672	07_BC
AX149771	CN54b	07_BC
AX149898	C54C	07_BC

AF286229	98CN006	08 BC
AY008715	97CNGX6F	08 BC
AY008716	97CNGX7F	08 BC
AY008717	97CNGX9F	08 BC
AF289548	96TZBF061	10 CD
AF289549	96TZBF071	10 CD
AF289550	96TZBF110	10 CD
AF179368	GR17	11 cpx
AJ291718	MP818	11 cpx
AJ291719	MP1298	11 cpx
AJ291720	MP1307	11 cpx
AF385934	URTR23	12 BF
AF385935	URTR35	12 BF
AF385936	ARMA159	12 BF
AF408629	A32879 AF408629	12 BF
AF408630	A32989 AF408630	12 BF
AY037279	ARMA185	12 BF
AF423756	X397 AF423756	14 BG
AF423757	X421 AF423757	14 BG
AF423758	X475 AF423758	14 BG
AF423759	X477 AF423759	14 BG
AF450096	X605 AF450096	14 BG
AF450097	X623 AF450097	14 BG
AF069669	SE8538	A
AF069671	SE7535	A
AF069673	SE8891	A
AF107771	UGSE8131	A
AF193275	97BL006 AF193275	A
AF361872	97TZ02 AF361872	A
AF361873	97TZ03 AF361873	A
AF413987	98UA0116 AF413987	A
AF004885	Q23-17	A1
AF069670	SE7253	A1
M62320	U455 U455A	A1
U51190	92UG037	A1
AF286237	94CY017.41	A2
AF286238	97CDKTB48	A2
U86780	ZAM184	A2C
AF286239	97KR004	A2D
AF316544	97CDKP58	A2G
AF067156	95IN21301	AC
AF071474	SE9488	AC
AF361871	97TZ01 AF361871	AC
AF361876	97TZ06 AF361876	AC
AF361878	97TZ08 AF361878	AC
AF361879	97TZ09 AF361879	AC
U88823	92RW009	AC
AF075702	SE8603	ACD
AJ276595	VI1035	ACG
AF071473	SE7108	AD
AF075701	SE6954	AD
AJ237565	97NOGIL3	ADHK
X04415	MAL MALCG	ADK
AF377959	CM53379 AF377959	AFGHJU
AF377957	CM53392 AF377957	AG



AJ276596	VI1197	AG
U88825	92NG003	AG
AF076474	VI354	AGHU
AF192135	BW2117	AGJ
AJ293865	B76 HIM293865	AGJ
AF069672	SE6594	AU
A04321	IIIB LAI	B
AB078005	ARES2 AB078005	B
AF003887	WC001	B
AF003888	NL43WC001	B
AF004394	AD87 ADA	B
AF033819	HXB2-copy LAI	B
AF042100	MBC200	B
AF042101	MBC925	B
AF042102	MBC18 MBCC18	B
AF042103	MBCC54	B
AF042104	MBCC98	B
AF042105	MBCD36	B
AF042106	MBCC18R01 C18R01	B
AF049494	499JC16	B
AF049495	NC7	B
AF069140	DH12-3	B
AF070521	NL43E9 LAI IIIB/NY5	B
AF075719	MNTQ MNcloneTQ	B
AF086817	TWCYS LM49	B
AF146728	VH	B
AF224507	WK	B
AF256204	S61I1 AF256204	B
AF256205	S61D15 AF256205	B
AF256206	S61G1 AF256206	B
AF256207	S61G7 AF256207	B
AF256208	S61I15 AF256208	B
AF256209	S61K1 AF256209	B
AF256210	S61K15 AF256210	B
AF256211	S61D1	B
AF286365	WR27	B
AJ006287	89SP061 89ES061	B
AJ271445	GB8 GB8-46R HIM271445	B
AX078307	BH10	B
AY037268	ARCH054	B
AY037269	ARMS008	B
AY037270	BOL122	B
AY037274	ARMA173	B
AY037282	ARMA132	B
D10112	CAM1	B
D86068	MCK1	B
D86069	PM213	B
K02007	SF2 LAV2 ARV2	B
K02013	LAI BRU	B
K02083	PV22	B
K03455	HXB2 HXB2CG HXB2R LAI	B
L02317	BC BCSG3	B
L31963	TH475A LAI	B
M15654	BH102 BH10	B
M17449	MNCG MN	B

M17451	RF HAT3	B
M19921	NL43 pNL43 NL4-3	B
M26727	OYI, 397	B
M38429	JRCSF JR-CSF	B
M38431	NY5CG	B
M93258	YU2 YU2X	B
M93259	YU10	B
NC 001802	HXB2R	B
U12055	LW123	B
U21135	WEAU160 GHOSH	B
U23487	contaminant MANC	B
U26546	WR27	B
U26942	NL4-3 LAI/NY5 pNL43 NL43	B
U34603	H0320-2A12 ACH3202A12	B
U34604	3202A21 ACH3202A21	B
U37270	C18MBC	B
U39362	P896 89.6	B
U43096	D31	B
U43141	HAN	B
U63632	JRFL JR-FL	B
U69584	85WCIPR54	B
U69585	WCIPR854	B
U69586	WCIPR8546	B
U69587	WCIPR8552	B
U69588	WCIPR855	B
U69589	WCIPR9011	B
U69590	WCIPR9012	B
U69591	WCIPR9018	B
U69592	WCIPR9031	B
U69593	WCIPR9032	B
U71182	RL42	B
X01762	REHTLV3 LAI IIIB	B
Z11530	F12CG	B
AF332867	A027 AF332867	BF
AF408626	A025 AF408626	BF
AF408627	A047 AF408627	BF
AF408628	A063 AF408628	BF
AF408631	A050 AF408631	BF
AF408632	A32878 AF408632	BF
AY037266	ARCH014	BF
AY037267	ARCH003	BF
AY037271	BOL137	BF
AY037272	URTR17	BF
AY037273	ARMA062	BF
AY037275	ARMA036	BF
AY037276	ARMA070	BF
AY037277	ARMA037	BF
AY037278	ARMA006	BF
AY037280	ARMA097	BF
AY037281	ARMA038	BF
AY037283	ARMA029	BF
AF005495	93BR029.4	BF1
AF423755	X254 AF423755	BG
AB023804	93IN101	C
AF067154	93IN999 301999	C

AF067155	95IN21068	C
AF067157	93IN904 301904	C
AF067158	93IN905 301905	C
AF067159	94IN11246	C
AF110959	96BW01B03 96BW01B03	C
AF110960	96BW01B21	C
AF110961	96BW01B22	C
AF110962	96BW0402	C
AF110963	96BW0407	C
AF110964	96BW0408	C
AF110965	96BW0409	C
AF110966	96BW0410	C
AF110967	96BW0502	C
AF110968	96BW0504	C
AF110969	96BW1104	C
AF110970	96BW1106	C
AF110971	96BW11B01	C
AF110972	96BW1210	C
AF110973	96BW15B03	C
AF110974	96BW15C02	C
AF110975	96BW15C05	C
AF110976	96BW16B01	C
AF110977	96BW16D14	C
AF110978	96BW1626	C
AF110979	96BW17A09	C
AF110980	96BW17B03	C
AF110981	96BW17B05	C
AF286223	94IN476	C
AF286224	96ZM651	C
AF286225	96ZM751	C
AF286227	97ZA012	C
AF286228	98BR004	C
AF286231	98IN012	C
AF286232	98IN022	C
AF286233	98IS002	C
AF286234	98TZ013	C
AF286235	98TZ017	C
AF290027	96BW06H51 96BW06-H51	C
AF290028	96BW06J4	C
AF290029	96BW06J7 AF290029	C
AF290030	96BW06K18 AF290030	C
AF321523	MJ4	C
AF361874	97TZ04 AF361874	C
AF361875	97TZ05 AF361875	C
AF443074	96BWM015	C
AF443075	96BWM032 AF443075	C
AF443076	98BWM0122 AF443076	C
AF443077	98BWM0134 AF443077	C
AF443078	98BWM014A3 AF443078	C
AF443079	98BWM01410 AF443079	C
AF443080	98BWM018D5 AF443080	C
AF443081	98BWM036A5 AF443081	C
AF443082	98BWM037D5 AF443082	C
AF443083	99BW393212 AF443083	C
AF443084	99BW46424 AF443084	C

AF443085	99BW47458 AF443085	C
AF443086	99BW47547 AF443086	C
AF443087	99BWMC168 AF443087	C
AF443088	00BW07621 AF443088	C
AF443089	00BW076820 AF443089	C
AF443090	00BW087421 AF443090	C
AF443091	00BW147127 AF443091	C
AF443092	00BW16162 AF443092	C
AF443093	00BW1686. 00BW16868 AF443093	C
AF443094	00BW17593 AF443094	C
AF443095	00BW17732 AF443095	C
AF443096	00BW17835 AF443096	C
AF443097	00BW17956 AF443097	C
AF443098	00BW18113 AF443098	C
AF443099	00BW18595 AF443099	C
AF443100	00BW18802 AF443100	C
AF443101	00BW192113 AF443101	C
AF443102	00BW20361 AF443102	C
AF443103	00BW20636 AF443103	C
AF443104	00BW20872 AF443104	C
AF443105	00BW2127214 AF443105	C
AF443106	00BW21283 AF443106	C
AF443107	00BW22767 AF443107	C
AF443108	00BW38193 AF443108	C
AF443109	00BW38428 AF443109	C
AF443110	00BW38713 AF443110	C
AF443111	00BW38769	C
AF443112	00BW38868	C
AF443113	00BW38916	C
AF443114	00BW39702	C
AF443115	00BW50311	C
AY043173	DU151 AY043173	C
AY043174	DU179 AY043174	C
AY043175	DU422 AY043175	C
AY043176	CTSC2 AY043176	C
U46016	ETH2220 C2220	C
U52953	92BR025	C
AF361877	97TZ07 AF361877	CD
AY074891	00BWMO351 AY074891	CD
AF133821	MB2059	D
AJ320484	HIM320484	D
K03454	ELI	D
M22639	Z2Z6 Z2 CDC-Z34	D
M27323	NDK	D
U88822	84ZR085	D
U88824	94UG1141	D
AF005494	93BR020.1	F1
AF075703	FIN9363	F1
AF077336	VI850	F1
AJ249238	MP411 96FRMP411	F1
AF377956	CM53657 AF377956	F2
AJ249236	MP255 95CMMP255	F2
AJ249237	MP257 95CM-MP257C	F2
AF076475	VI1126	F2KU
AF061640	HH8793-1.1	G

AF061641	HH8793-12.1	G
AF061642	SE6165 G6165	G
AF084936	DRCBL	G
AF423760	X558 AF423760	G
AF450098	X138 AF450098	G
U88826	92NG083 JV10832	G
AF005496	90CF056 90CR056	H
AF190127	VI991	H
AF190128	VI997	H
AF082394	SE7887 SE92809	J
AF082395	SE7022 SE9173	J
AJ249235	EQTB11C 97ZR-EQTB11C	K
AJ249239	MP535 96CM-MP535C	K
AJ239083	97CA-MP645M/O	MO
AJ006022	YBF30	N
AJ271370	YBF106	N
AF407418	VAU AF407418	O
AF407419	VAU AF407419	O
AJ302646	SEMP1299 HIM302646	O
AJ302647	SEMP1300 HIM302647	O
L20571	MVP5180	O
L20587	ANT70	O
NC 002787	SEMP1299 NC 002787	O
AF286236	83CD003 Z3 AF286236	U
AF457101	90CD121E12 AF457101	U
AY046058	GR303 99GR303 AY046058	U

(700000000)

Table II: HIV siNA and Target Sequences

Target Sequence	SeqID	Upper seq	Seq ID	Lower seq	Seq ID
UUUGGAAAGGACAGCAAA	1	UUUGGAAAGGACAGCAAA	1	UUUGCUGGCUUUUCCAAA	739
CAGGACAGAUAGAUACAGU	2	CAGGACAGAUAGAUACAGU	2	ACUGUAUCAUCUGCUCCUG	740
AGAAAAGGGGGAUUGGGG	3	AGAAAAGGGGGAUUGGGG	3	CCCCAAUCCCCCUUUUCU	741
GUAGACAGGAUGAGGAUUA	4	GUAGACAGGAUGAGGAUUA	4	UAAUCCUAUCCUGUCUAC	742
ACAGGAGCAGAUAGAUACAG	5	ACAGGAGCAGAUAGAUACAG	5	CUGUAUCAUCUGCUCCUGU	743
GAAAAGGGGGAUUGGGG	6	GAAAAGGGGGAUUGGGG	6	CCCCAAUCCCCCUUUUC	744
UUAGAUACAGGAGCAGAU	7	UUAGAUACAGGAGCAGAU	7	CAUCUGCUCCUGUAUCUAA	745
UAGAUACAGGAGCAGAU	8	UAGAUACAGGAGCAGAU	8	UCAUCUGCUCCUGUAUCUA	746
AGCAGAAACAGUGGCAAU	9	AGCAGAAACAGUGGCAAU	9	AUUGCCACUGUCUUCUGCU	747
AUUAGAUACAGGAGCAGAU	10	AUUAGAUACAGGAGCAGAU	10	AUCUGCUCCUGUAUCUAAU	748
AUACAGGAGCAGAUAC	11	AUACAGGAGCAGAUAC	11	GUUAUCAUCUGCUCCUGUAU	749
GAGCAGAAGACAGUGGCAA	12	GAGCAGAAGACAGUGGCAA	12	UUGCCACUGUCUUCUGCUC	750
AGAGCAGAAGACAGUGGCA	13	AGAGCAGAAGACAGUGGCA	13	UGCCACUGUCUUCUGCUCU	751
GCAGAAGACAGUGGCAUUG	14	GCAGAAGACAGUGGCAUUG	14	CAUUGCCACUGUCUUCUGC	752
AGAUACAGGAGCAGAU	15	AGAUACAGGAGCAGAU	15	AUCAUCUGCUCCUGUAUCU	753
UACAGGAGCAGAUAGAUACA	16	UACAGGAGCAGAUAGAUACA	16	UGUAUCAUCUGCUCCUGUA	754
UAUUAGAUACAGGAGCAGA	17	UAUUAGAUACAGGAGCAGA	17	UCUGCUCCUGUAUCUAAUA	755
GAUACAGGAGCAGAUAGUA	18	GAUACAGGAGCAGAUAGUA	18	UAUCAUCUGCUCCUGUAUC	756
AUGGAAACAGAUUGGCAGG	19	AUGGAAACAGAUUGGCAGG	19	CCUGCCAUUCUUUUCCAU	757
GUCAACAUAAUUGGAAGAA	20	GUCAACAUAAUUGGAAGAA	20	UUCUCCAAUUAUGUUGAC	758
UAUGGAAACAGAUUGGCAG	21	UAUGGAAACAGAUUGGCAG	21	CUGCCAUUCUUUUCCAU	759
AUGAUAGGGGGAUUGGAG	22	AUGAUAGGGGGAUUGGAG	22	CUCCAAUUCUUUUCCAU	760
CAGAAGACAGUGGCAUUGA	23	CAGAAGACAGUGGCAUUGA	23	UCAUUGCCACUGUCUUCUG	761
CAUUGGCCAUUUGACAGAAG	24	CAUUGGCCAUUUGACAGAAG	24	CUUCUGUCAAUGGCCAUUG	762
UCAACAUAAUUGGAAGAAA	25	UCAACAUAAUUGGAAGAAA	25	UUUCUCCAAUUAUGUUGA	763
AAUGGCCAUUUGACAGAAGA	26	AAUGGCCAUUUGACAGAAGA	26	UCUUCUGUCAUUGGCCAUU	764
UGAUAGGGGGAUUGGAGG	27	UGAUAGGGGGAUUGGAGG	27	CCUCCAAUUCUUUUCCAU	765
GACAGGCUAUUUUUAGG	28	GACAGGCUAUUUUUAGG	28	CCUAAAAUUUAGCCUGUC	766
AUUUCGGGUUUAUACAG	29	AUUUCGGGUUUAUACAG	29	CUGUAUUAACCCGAAAAU	767
CUAUUAGAUACAGGAGCAG	30	CUAUUAGAUACAGGAGCAG	30	CUGCUCCUGUAUCUAAUAG	768

AGACAGGCUAUUUUUUUAG	31	AGACAGGCUAUUUUUUUAG	31	CUAAAAUUUAGCCUGUCU	769
AAUUGAUAGGGGAAUUGG	32	AAUUGAUAGGGGAAUUGG	32	CCAAUUCUUUUUUAUUAUU	770
UAUGGGCAAGCAGGGAGCU	33	UAUGGGCAAGCAGGGAGCU	33	AGCUCCUGCUUUGCCCAUA	771
UAGUAGGGCAAGCAGGGA	34	UAGUAGGGCAAGCAGGGA	34	UCCUGCUUUGCCCAUACUA	772
GAAACAGAUUGCAGGUGA	35	GAAACAGAUUGCAGGUGA	35	UCACUGCCCAUCUGUUUUC	773
ACCAUCAUAGGGAAGCUG	36	ACCAUCAUAGGGAAGCUG	36	CAGCUUCCCAUUGAUGGU	774
AAUGAUAGGGGAAUUGGA	37	AAUGAUAGGGGAAUUGGA	37	UCCAAUUCUUUUUUAUUA	775
UGGAAACAGAUUGCAGGU	38	UGGAAACAGAUUGCAGGU	38	ACUGCCCAUCUGUUUCCA	776
GGAAACAGAUUGCAGGUG	39	GGAAACAGAUUGCAGGUG	39	CACCUGCCAUCUGUUUUC	777
GAUUAUGGAAACAGAUUG	40	GAUUAUGGAAACAGAUUG	40	CCAUCUGUUUUUCCAAUUC	778
AAAUGAUAGGGGAAUUG	41	AAAUGAUAGGGGAAUUG	41	CAAUUCUUUUUUAUUAUU	779
UGGAAAGGUGAAGGGGCAG	42	UGGAAAGGUGAAGGGGCAG	42	CUGCCCUUCCACCUUCCA	780
AUCAUAGGGAAGCUGCAG	43	AUCAUAGGGAAGCUGCAG	43	CUGCAGCUUCCCAUUAU	781
UGGAAACCAAAAAUAGUAG	44	UGGAAACCAAAAAUAGUAG	44	CUAUCAUUUUUUGUUUCCA	782
CCAUCAUAGGGAAGCUGC	45	CCAUCAUAGGGAAGCUGC	45	GCAGCUUCCCAUUAUGG	783
AGGGAUUAUGGAAACAGA	46	AGGGAUUAUGGAAACAGA	46	UCUGUUUCCCAUUAUCCU	784
GGAAACCAAAAAUAGUAGG	47	GGAAACCAAAAAUAGUAGG	47	CCUAUCAUUUUUGUUUCC	785
UAGGGGCAUUGGAGGUUU	48	UAGGGGCAUUGGAGGUUU	48	AAACCUCAAUUCCCUUA	786
UACAGUGCAGGGGAAAGAA	49	UACAGUGCAGGGGAAAGAA	49	UUCUUUUUUUCCGACUGUA	787
CUCUAUAGAUACAGGAGC	50	CUCUAUAGAUACAGGAGC	50	GCUCUGUAUUAUUAUAG	788
GGAUUAUGGAAACAGAUUG	51	GGAUUAUGGAAACAGAUUG	51	CAUCUGUUUUUCCAAUUC	789
CCAAAAUAGUAGGGGGAA	52	CCAAAAUAGUAGGGGGAA	52	UUCUUUUUUUUAUUAUUG	790
AUGGAAACCAAAAAUAGUA	53	AUGGAAACCAAAAAUAGUA	53	UAUCAUUUUUUUGUUUCCA	791
CAGUGCAGGGGAAAGAAUA	54	CAGUGCAGGGGAAAGAAUA	54	UAUUCUUUUUUUUUCCUG	792
ACAAUGGCCAUUGACAGAA	55	ACAAUGGCCAUUGACAGAA	55	UUCUGUCAUUGGCCAUUGU	793
CCAUGCAUGGACAAGUAGA	56	CCAUGCAUGGACAAGUAGA	56	UCUACUUUGUCCCAUGG	794
AUUAUGGAAACAGAUUGC	57	AUUAUGGAAACAGAUUGC	57	GCCAUUCUUUUUCCAAUUA	795
AACAUGGCCAUUGACAGA	58	AACAUGGCCAUUGACAGA	58	UCUGUCAUUGGCCAUUGUU	796
AAAAUGAUAGGGGAAUUA	59	AAAAUGAUAGGGGAAUUA	59	AAUUCUUUUUUAUUAUUU	797
GCCAUUGGACAAGUAG	60	GCCAUUGGACAAGUAG	60	CUACUUGUCCCAUGG	798
UAGCAGGAAGAUUGCCAGU	61	UAGCAGGAAGAUUGCCAGU	61	ACUGGCCAUUCUCCUGUA	799
CAAAAAUAGUAGGGGAAU	62	CAAAAAUAGUAGGGGAAU	62	AUUCUUUUUUAUUAUUG	800
AAGAAUAGUAGACAGAUUG	63	AAGAAUAGUAGACAGAUUG	63	CAUGCUGUCAUUAUUAU	801
UCUAUUAUAGACAGGCA	64	UCUAUUAUAGACAGGCA	64	UGCUCUGUAUUAUUAUAG	802
GCUCUAUUAUAGACAGGAG	65	GCUCUAUUAUAGACAGGAG	65	CUCUCUGUAUUAUUAUAGC	803
CAGGCUAAUUUUUUUAGGGA	66	CAGGCUAAUUUUUUUAGGGA	66	UCCCUAAAAUUUAGCCUG	804

(TUUUUUU)

AGGAGCAGAUUAUACAGUA	67	AGGAGCAGAUUAUACAGUA	67	UACUGUAUUAUACUGCUCCU	805
AAACAAUUGGCCAUUAGACAG	68	AAACAAUUGGCCAUUAGACAG	68	CUGUCAUUGGCCAUUUGUUU	806
CGGGUUUUUAUACAGGACA	69	CGGGUUUUUAUACAGGACA	69	UGUCCUGUAUUAUAAACCCG	807
CAACAUAAUUGGAAGAAU	70	CAACAUAAUUGGAAGAAU	70	AUUUUAUCCAAUUUAUUGUUG	808
UCAAUAGGAAGCUGCAGA	71	UCAAUAGGAAGCUGCAGA	71	UCUGCAGCUUCCUUAUUGA	809
GGAAAGGUUAGGGGAGU	72	GGAAAGGUUAGGGGAGU	72	ACUGCCCUUACCUUUUC	810
UUUCGGUUUAUUAACAGGG	73	UUUCGGUUUAUUAACAGGG	73	CCUGUAUUAUAAACCCGAA	811
UCGGUUUAUUAACAGGGAC	74	UCGGUUUAUUAACAGGGAC	74	GUCCUGUAUUAUAAACCCGA	812
ACAGUGCAGGGGAAAGAAU	75	ACAGUGCAGGGGAAAGAAU	75	AUUCUUUCCCUUGGCACUGU	813
AUGCAUGGACAAGUAGACU	76	AUGCAUGGACAAGUAGACU	76	AGUCUACUUUGUCCAUUGCAU	814
AAGCCAUGCAUGGACAAGU	77	AAGCCAUGCAUGGACAAGU	77	ACUUUGUCCAUUGCAUGGCUU	815
AGCCAUGCAUGGACAAGUA	78	AGCCAUGCAUGGACAAGUA	78	UACUUUGUCCAUUGCAUGGCU	816
GCAUUUAUCAGAGGAGCCA	79	GCAUUUAUCAGAGGAGCCA	79	UGGCUCUUCUUGAUUAUGC	817
AAUUGGAGAGGAGUUAU	80	AAUUGGAGAGGAGUUAU	80	AUAUUUACAUUCUCCAAUU	818
AGAAAAUUCAGUAACAGU	81	AGAAAAUUCAGUAACAGU	81	ACUGUUACUGAUUUUUUUUCU	819
GAAGCCAUGCAUGGACAAG	82	GAAGCCAUGCAUGGACAAG	82	CUUGUCCAUUGCAUGGCUUC	820
ACAGGCUAAUUUUUAGGG	83	ACAGGCUAAUUUUUAGGG	83	CCCUAAAAAUUAGCCUGU	821
GAAGAAUUAUGACAGCAU	84	GAAGAAUUAUGACAGCAU	84	AGCUGUCAUUAUUUCUUC	822
UUUUCGGUUUAUUAACAGG	85	UUUUCGGUUUAUUAACAGG	85	CCUGUAUUAUAAACCCGAAA	823
ACCAAAAUUAUAGGGGGA	86	ACCAAAAUUAUAGGGGGA	86	UCCCCUUAUUAUUUUGGU	824
GAAGUGACAUAAGGGAAC	87	GAAGUGACAUAAGGGAAC	87	GUUCCUGCUAUGUCACUUC	825
UUCGGUUUAUUAACAGGGA	88	UUCGGUUUAUUAACAGGGA	88	UCCCUUAUUAUAAACCCGAA	826
AUAGGGGGAUUUGGAGGUU	89	AUAGGGGGAUUUGGAGGUU	89	AACCUCCAUIUCCCCCUAU	827
AGAAGAAUUAUGACAGCA	90	AGAAGAAUUAUGACAGCA	90	UGCUGUCAUUAUUUGUUCU	828
AUUGGAGAAGUGAAUUAUA	91	AUUGGAGAAGUGAAUUAUA	91	UAUAUUUACAUUCUCGCAAU	829
GGAAUGACAUAGCAGGAA	92	GGAAUGACAUAGCAGGAA	92	UUCUGCUAUGUCACUUC	830
AGGCUAAUUUUUAGGGAA	93	AGGCUAAUUUUUAGGGAA	93	UUCCCUAAAAAUUAGCCU	831
UUUUGGAAAACAGAU	94	UUUUGGAAAACAGAU	94	UGCCAUCUGUUUUUCCAUAA	832
GGGAUUAUGGAAAACAGAU	95	GGGAUUAUGGAAAACAGAU	95	AUCUGUUUUCCAUAAUCCC	833
UAGAAGAAUUAUGACAGC	96	UAGAAGAAUUAUGACAGC	96	GCUGUCAUUAUUUCUUA	834
AGCUCUUAUUAUUAACAGGA	97	AGCUCUUAUUAUUAACAGGA	97	UCCUGUAUUAUUAUAGAGCU	835
GUUUGGGAAGCAGGGGAGC	98	GUUUGGGAAGCAGGGGAGC	98	GCUCUCCUGUUGCCCAUAC	836
CUUAGGCAUCUCCUUAUGGC	99	CUUAGGCAUCUCCUUAUGGC	99	GCCAUAGGAGAUCCCUAAG	837
GCAGGAACUACUAGUACCC	100	GCAGGAACUACUAGUACCC	100	GGGUACUAGUAGUUCUUGC	838
GGGGAAGUGACAUAGCAGG	101	GGGGAAGUGACAUAGCAGG	101	CCUGCUAUGUCACUUCCCC	839
UACAAUCCCCAAAGUCAAG	102	UACAAUCCCCAAAGUCAAG	102	CUUGACUUUUGGGGAUUGUA	840



U U C C C U A C A A U C C C C A A A G	103	U U C C C U A C A A U C C C C A A A G	103	C U U G G G G A U U G U A G G G A A	841
A A G C U C U A U U A G A U A C A G G	104	A A G C U C U A U U A G A U A C A G G	104	C C U G U A U C U A A U A G A G C U U	842
C C U A U G G C A G G A A G A G C G	105	C C U A U G G C A G G A A G A G C G	105	C G C U U C U U C C U G C C A U A G G	843
A G G G G A A G U G A C A U A G C A G	106	A G G G A A G U G A C A U A G C A G	106	C U G C U A U G U C A C U U C C C C U	844
U C C U A U G G C A G G A A G A G C	107	U C C U A U G G C A G G A A G A G C	107	G C U U C U U C C U G C C A U A G G A	845
C A G C A U U A U C A G A A G A G C	108	C A G C A U U A U C A G A A G A G C	108	G C U C C U U C U G A U A A U G C U G	846
A U C U C C U A U G G C A G G A A G A	109	A U C U C C U A U G G C A G G A A G A	109	U C U U C C U G C C A U A G G A G A U	847
A G C A G G A A C U A C U A G U A C C	110	A G C A G G A A C U A C U A G U A C C	110	G G U A C U A G U A G U U C C U G C U	848
G A A C C A A A A A U G A U A G G G	111	G A A C C A A A A A U G A U A G G G	111	C C C U A C A U U U U U G G U U U C	849
A A C C A A A A A U G A U A G G G G	112	A A C C A A A A A U G A U A G G G G	112	C C C C U A C A U U U U U G G U U U	850
C A G A A G G A G C C A C C C C A C A	113	C A G A A G G A G C C A C C C C A C A	113	U G U G G G G U G G C U C C U U C U G	851
U A G C A G G A A C U A C U A G U A C	114	U A G C A G G A A C U A C U A G U A C	114	G U A C U A G U A G U U C C U G C U A	852
U G C A U G G A C A A G U A G A C U G	115	U G C A U G G A C A A G U A G A C U G	115	C A G U C U A C U U G U C C A U G C A	853
U U A G G C A U C U C C U A U G G C A	116	U U A G G C A U C U C C U A U G G C A	116	U G C C A U A G G A G A U G C C U A A	854
U A U G G C A G G A A G A G C G G A	117	U A U G G C A G G A A G A G C G G A	117	U C C G C U U C U U C C U G C C A U A	855
A U A G C A G G A A C U A C U A G U A	118	A U A G C A G G A A C U A C U A G U A	118	U A C U A G U A G U U C C U G C U A U	856
U A G A C A U A A U A G C A A C A G A	119	U A G A C A U A A U A G C A A C A G A	119	U C U G U U G C U A U U A U A U G U C U A	857
C A U U A U C A A G G A G C C C A C	120	C A U U A U C A A G G A G C C C A C	120	G U G C U C C U U C U G A U A A U G	858
C U A U G G C A G A A G A A G C G G	121	C U A U G G C A G A A G A A G C G G	121	C G C U U C U U C C U G C C A U A G	859
G A U A G G G G A A U U G G A G G U	122	G A U A G G G G A A U U G G A G G U	122	A C C U C C A A U U C C C C C U A U C	860
A C A A U C C C C A A A G U C A A G G	123	A C A A U C C C C A A A G U C A A G G	123	C C U U G A C U U U G G G A U U G U	861
A U U C C C U A C A A U C C C C A A A	124	A U U C C C U A C A A U C C C C A A A	124	U U U G G G A U U G U A G G G A A U	862
A A C C A A A A A U G A U A G G G G G	125	A A C C A A A A A U G A U A G G G G G	125	C C C C C U A C A U U U U U G G U U	863
U C U C C U A U G G C A G G A A G A A	126	U C U C C U A U G G C A G G A A G A A	126	U U C U U C C U G C C A U A G G A G A	864
C A U G C A U G G A C A A G U A G A C	127	C A U G C A U G G A C A A G U A G A C	127	G U C U A C U U G U C C A U A G C A U G	865
C C U G U A C C C A C A G A C C C	128	C C U G U A C C C A C A G A C C C	128	G G G U C U G U G G G U A C A C A G G	866
C A U C A A U G A G G A A G C U G C A	129	C A U C A A U G A G G A A G C U G C A	129	U G C A G C U U C C U C A U U G A U G	867
G A C A U A G C A G G A A C U A C U A	130	G A C A U A G C A G G A A C U A C U A	130	U A G U A G U U C C U G C U A U G U C	868
G A A A G G U A A G G G G C A G U A	131	G A A A G G U A A G G G G C A G U A	131	U A C U G C C C U U C A C C U U U C	869
A G U G A C A U A G C A G G A A C U A	132	A G U G A C A U A G C A G G A A C U A	132	U A G U U C C U G C U A U G U C A C U	870
G C A G A U G A U A C A G A U U A G	133	G C A G A U G A U A C A G A U U A G	133	C U A A U A C U G U A U C A U C U G C	871
G G A G C A G A U G A U A C A G U A U	134	G G A G C A G A U G A U A C A G U A U	134	A U A C U G U A U C A U C U G C U C C	872
C C A A G G G A A G U G A C A U A G	135	C C A A G G G A A G U G A C A U A G	135	C U A U G U C A C U U C C C U U G G	873
G A A G C U C U A U U A G A U A C A G	136	G A A G C U C U A U U A G A U A C A G	136	C U G U A U C A A U A G A G C U U C	874
G G G A A G U G A C A U A G C A G G A	137	G G G A A G U G A C A U A G C A G G A	137	U C C U G C U A U G U C A C U U C C C	875
C A U G C C U G U G A C C C A C A G	138	C A U G C C U G U G A C C C A C A G	138	C U G U G G G U A C A C A G G C A U G	876

GAAAGAGCAGAAAGACAGUG	139	GAAAGAGCAGAAAGACAGUG	139	CACUGUCUUUCUGCUCUUUC	877
ACAUAGCAGGAACUACUAG	140	ACAUAGCAGGAACUACUAG	140	CUAGUAGUUCCUGCUAUAGU	878
CAUCUCCUUAUGGCGGAAG	141	CAUCUCCUUAUGGCGGAAG	141	CUUCCUGCCAUAGGAGAU	879
GAGCAGAUAGUACAGUAAU	142	GAGCAGAUAGUACAGUAAU	142	AAUACUGUAUCAUCUGCUC	880
AGCAUUAUCAGAGGAGCC	143	AGCAUUAUCAGAGGAGCC	143	GGCUCCUUCUGUAUAAUGCU	881
CACCAGGCCAGAUAGAGAGA	144	CACCAGGCCAGAUAGAGAGA	144	UCUCUCAUCUGGCCUGGUG	882
GUGACAUAGCAGGAACUAC	145	GUGACAUAGCAGGAACUAC	145	GUAGUCCUGCUAUUGUCAC	883
AGCAGGAAGAUAGGCGAGUA	146	AGCAGGAAGAUAGGCGAGUA	146	UACUGGCCAUUCUCCUGCU	884
GAGAACCAAGGGGAAGUGA	147	GAGAACCAAGGGGAAGUGA	147	UCACUUCUCCUUGGUUUCUC	885
AGUAUGGCAAGCAGGGAG	148	AGUAUGGCAAGCAGGGAG	148	CUCCUGCUUGGCCAUACU	886
CCUACAUCUCCCAAGUCA	149	CCUACAUCUCCCAAGUCA	149	UGACUUCUUGGGGAUUGUAG	887
CUACAUCUCCCAAGUCA	150	CUACAUCUCCCAAGUCA	150	UUGACUUCUUGGGGAUUGUAG	888
GCCUGUACCCACAGACC	151	GCCUGUACCCACAGACC	151	GGUCUGUGGGGUACACAGGC	889
AGCAGUAGUACAGUAAUUA	152	AGCAGUAGUACAGUAAUUA	152	UAAUACUGUAUCAUCUGCU	890
AGAGAACCCAGGGAGUG	153	AGAGAACCCAGGGAGUG	153	CACUUCUCCUUGGUUCUCU	891
CCCUACAUCUCCCAAGUC	154	CCCUACAUCUCCCAAGUC	154	GACUUCUUGGGGAUUGUAGGG	892
UGACAUAGCAGGAACUACU	155	UGACAUAGCAGGAACUACU	155	AGUAGUCCUGCUAUUGUCA	893
UUAUCAGAAAGGAGCCACCC	156	UUAUCAGAAAGGAGCCACCC	156	GGUGGCUCCUUCUGAUAA	894
AAGUGACAUAGCAGGAACU	157	AAGUGACAUAGCAGGAACU	157	AGUUCUCCUGCUAUUGUCACUU	895
GCAGGAAGAUAGGCGAGUAA	158	GCAGGAAGAUAGGCGAGUAA	158	UUAUCUGGCCAUUCUCCUGC	896
UAGGCAUCUCCUUAUGGCAG	159	UAGGCAUCUCCUUAUGGCAG	159	CUGCCAUAGGAGAUGCCUA	897
CAAGGGGAAGUGACAUAGC	160	CAAGGGGAAGUGACAUAGC	160	GCUAUGUCACUUCUCCUUG	898
AAAGAGCAGAAAGACAGUGG	161	AAAGAGCAGAAAGACAGUGG	161	CCACUGUCUUCUGCUCUUU	899
CUCCUUAUGGCGAGGAAG	162	CUCCUUAUGGCGAGGAAG	162	CUUCUUCUCCUGCCAUAGGAG	900
UAUCAGAAAGGAGCCACCC	163	UAUCAGAAAGGAGCCACCC	163	GGGUGGCUCCUUCUGAUUA	901
AUUAUCAGAAAGGAGCCACC	164	AUUAUCAGAAAGGAGCCACC	164	GGUGGCUCCUUCUGAUAAU	902
AUGCCUUGUAGCCACAGA	165	AUGCCUUGUAGCCACAGA	165	UCUGUGGCUACACAGGCAU	903
AAUUAAGUAGAUUUCAGAG	166	AAUUAAGUAGAUUUCAGAG	166	CUCUGAAUUCUACUAAUUU	904
UGCAUUAAGCAGCUGCUU	167	UGCAUUAAGCAGCUGCUU	167	AAGCAGCUGCUUUAUUGCA	905
AAUUAAGUAGAUUUCAGAGA	168	AAUUAAGUAGAUUUCAGAGA	168	UCUCUGAAUUCUACUAAUU	906
GCAUCUCUUAUGGCGAGAA	169	GCAUCUCUUAUGGCGAGAA	169	UUCUUGCCAUAGGAGUCC	907
AGAACCAAGGGGAAGUGAC	170	AGAACCAAGGGGAAGUGAC	170	GUCACUUCUCCUUGGUUCU	908
UCAAAAUUUCGGGUUUUAU	171	UCAAAAUUUCGGGUUUUAU	171	AUAAACCCGAAAAUUUUGA	909
CAGGGAUGGAAAGGAUCAC	172	CAGGGAUGGAAAGGAUCAC	172	GUGAUCUUCUCCAUCCUUG	910
GAGGAGGACCCCAAG	173	GAGGAGGACCCCAAG	173	CUUGGGGGUGGCUCCUUC	911
AAUUUUCGGGUUUUAUACA	174	AAUUUUCGGGUUUUAUACA	174	UGUAAUAAACCCGAAAAUU	912

(11111111)

AGCAGGAAGCACUAUGGGC	175	AGCAGGAAGCACUAUGGGC	175	GCCCAUAGUGCUUCCUGCU	913
AUCAGAAGGAGCCACCCCA	176	AUCAGAAGGAGCCACCCCA	176	UGGGUGGCUCCUUCUGAU	914
UGAGAGAACCAGGGGAAG	177	UGAGAGAACCAGGGGAAG	177	CUUCCCUUGGUUCUCUCA	915
AAGGUGAAGGGGCAGUAGU	178	AAGGUGAAGGGGCAGUAGU	178	ACUACUGCCCCUUCACCCU	916
GAAGAAUACAGUAACAGUA	179	GAAGAAUACAGUAACAGUA	179	UACUGUUACUGAUUUUUUG	917
CAUUGAGGAAGCUGCAGAA	180	CAUUGAGGAAGCUGCAGAA	180	UUCUGCAGCUUCCUCAUUG	918
AGAUGAUACAGUAUAGAA	181	AGAUGAUACAGUAUAGAA	181	UUCUAAUACUGUAUCAUCU	919
UGAGGAAGCUGCAGAAUGG	182	UGAGGAAGCUGCAGAAUGG	182	CCAUCUGCAGCUUCCUCA	920
UAUUUAGACCCAUCAAAG	183	UAUUUAGACCCAUCAAAG	183	CUUUUGAUGGGUCAAUAUA	921
UCACUCUUUGGCAACGACC	184	UCACUCUUUGGCAACGACC	184	GGUCGUUGCCAAAGAGUGA	922
UGGAGAAAUAUAGUAGAUU	185	UGGAGAAAUAUAGUAGAUU	185	AAUCUACUAAUUUUCUCCA	923
AGACAGGAUGAGGAUUAAGA	186	AGACAGGAUGAGGAUUAAGA	186	UCUAAUCCUCAUCCUGUCU	924
AAAGGUGAAGGGGCAGUAG	187	AAAGGUGAAGGGGCAGUAG	187	CUACUGCCCCUUCACCCUUU	925
GGCAUCUCUUAUGGCAGGA	188	GGCAUCUCUUAUGGCAGGA	188	UCCUGCCAUAGGAGAUGCC	926
AAGGAGCCACCCCAACAAGA	189	AAGGAGCCACCCCAACAAGA	189	UCUUGUGGGUGGCUCCUU	927
UAAAGCCAGGAUUGGAUGG	190	UAAAGCCAGGAUUGGAUGG	190	CCAUCCAUCCUGGCUUUA	928
GGAGAAAUAUAGUAGAUU	191	GGAGAAAUAUAGUAGAUU	191	AAAUACUAAUUUUCUCC	929
AAGAGCAGAGACAGUGGC	192	AAGAGCAGAGACAGUGGC	192	GCCACUGUCUUCUGCUCUU	930
UCAGAAGGAGCCACCCAC	193	UCAGAAGGAGCCACCCAC	193	GUGGGUGGCUCCUUCUGA	931
AGGAUCUCCUAUGGCAGG	194	AGGAUCUCCUAUGGCAGG	194	CCUGCCAUAGGAGAUGCCU	932
AGGAUGGAAAGAUCCACC	195	AGGAUGGAAAGAUCCACC	195	GGUGAUCCUUUCCAUCCCU	933
AGGAAGCUGCAGAAUGGGA	196	AGGAAGCUGCAGAAUGGGA	196	UCCCAUUCUGCAGCUUCCU	934
CUGCAUUAAGCAGCUGCU	197	CUGCAUUAAGCAGCUGCU	197	AGCAGCUGCUUUAUAGCAG	935
AAGGGGCAGUAGUAUACA	198	AAGGGGCAGUAGUAUACA	198	UGUAUUACUACUGCCCCUU	936
UUGACUAGCGGAGGCUAGA	199	UUGACUAGCGGAGGCUAGA	199	UCUAGCCUCCGCUAGUCA	937
UAAAAGACACCAAGGAAGC	200	UAAAAGACACCAAGGAAGC	200	GCUUCCUUGGUUGCUUUUA	938
GAGGAAGCUGCAGAAUGGG	201	GAGGAAGCUGCAGAAUGGG	201	CCCAUUCUGCAGCUUCCUC	939
CAGCAGGAAGCAGUAUUGG	202	CAGCAGGAAGCAGUAUUGG	202	CCCAUAGUGCUUCCUGCUG	940
GGAGCCACCCACAAGAUU	203	GGAGCCACCCACAAGAUU	203	AAUCUUGUGGGUGGCUCC	941
AUUAUGACCCACAAGAA	204	AUUAUGACCCACAAGAA	204	UCUUUUGAUGGGUCAUAU	942
CAGAUAUACAGUAUUAAGA	205	CAGAUAUACAGUAUUAAGA	205	UCUAAUACUGUAUCAUCUG	943
AUGAGAGAACCAGGGGAA	206	AUGAGAGAACCAGGGGAA	206	UCCCCUUGGUUCUCUCAU	944
AUGAGGAAGCUGCAGAAUG	207	AUGAGGAAGCUGCAGAAUG	207	CAUUCUGCAGCUUCCUCAU	945
UGCCUGUGUACCCACAGAC	208	UGCCUGUGUACCCACAGAC	208	GUCUGUGGUAACACAGGCA	946
GAAGGGGCAGUAGUAUAC	209	GAAGGGGCAGUAGUAUAC	209	GUUUUACUACUGCCCCUUC	947
UCAGCAUUUAUCAGAAGGAG	210	UCAGCAUUUAUCAGAAGGAG	210	CUCUUCUGUAUUAUGCUGA	948

UUCAAAAUUUCGGGUUUA	211	UUCAAAAUUUCGGGUUUA	211	UAAACCCGAAAAUUUUGAA	949
UCUGGAAAGGUGAAGGGC	212	UCUGGAAAGGUGAAGGGC	212	GCCCCUUCACCUUUCCAGA	950
UUAGCAGGAAGUAGGCCAG	213	UUAGCAGGAAGUAGGCCAG	213	CUGGCCAUUUCUCCUGCUAA	951
GAACCAAGGGGAAGUGACA	214	GAACCAAGGGGAAGUGACA	214	UGUCACUUCUCCUUGGUUC	952
AGAAGGAGCCACCCACAA	215	AGAAGGAGCCACCCACAA	215	UUUGUGGGUGGCUCCUUCU	953
AUUGAGGAAGCUGCAGAAU	216	AUUGAGGAAGCUGCAGAAU	216	AUUCUGCAGCUUCCUCAU	954
AAGAAAAUUCAGUAACAG	217	AAGAAAAUUCAGUAACAG	217	CUGUUAUGAUUUUUUUCUU	955
GGAAUUGGAGGUUUUAUCA	218	GGAAUUGGAGGUUUUAUCA	218	UGAUAAAAACCUCCAAUCC	956
UACAGUAUUAAGUAGGACCU	219	UACAGUAUUAAGUAGGACCU	219	AGGUCCUACUAUUAACUGUA	957
CCAGGAUUGGAGGCCCAA	220	CCAGGAUUGGAGGCCCAA	220	UUGGGCCAUCCAUUCCUGG	958
UUCUAUGUAGUAGGGGCAG	221	UUCUAUGUAGUAGGGGCAG	221	CUGCCCCAUUCACAUAGAA	959
CAAAAAUUUCGGGUUUUAU	222	CAAAAAUUUCGGGUUUUAU	222	AUAACCCGAAAAUUUUG	960
UAGACAGGAUGAGGAUUAG	223	UAGACAGGAUGAGGAUUAG	223	CUAAUCCUACUCCUGUCUA	961
UGACAGAAAGAAAAUAAA	224	UGACAGAAAGAAAAUAAA	224	UUUUAUUUUUUCUUCUGUCA	962
UUUAUUAACAGGACAGCAG	225	UUUAUUAACAGGACAGCAG	225	CUGCUUCCUCCUGUAUAAA	963
GGGUUUUAUACAGGGACAG	226	GGGUUUUAUACAGGGACAG	226	CUGUCCUCCUGUAUAAACC	964
AGAUUGAACAGCCCCAGA	227	AGAUUGAACAGCCCCAGA	227	UCUGGGGUUUUUCUCCAUUC	965
CUAGCGGAGGCUAGAAGGA	228	CUAGCGGAGGCUAGAAGGA	228	UCCUUCUAGCCUCCGCUAG	966
UGACUAGCGGAGGCUAGAA	229	UGACUAGCGGAGGCUAGAA	229	UUCUAGCCUCCGCUAGUCA	967
GACAUAAUAGCAACAGACA	230	GACAUAAUAGCAACAGACA	230	UGUCUGUUGCUAUUAUGUC	968
GGUUUAUUAACAGGGACAGC	231	GGUUUAUUAACAGGGACAGC	231	GCUGUCCUUGUAUAAACC	969
GCAGGUGAUGAUUGUGUGG	232	GCAGGUGAUGAUUGUGUGG	232	CCACACAUAUACACCUGC	970
AUGGCAGGAAGAGCGGAG	233	AUGGCAGGAAGAGCGGAG	233	CUCCGCUUCUCCUGCCAU	971
AGGUGAUGAUUGUGUGGCA	234	AGGUGAUGAUUGUGUGGCA	234	UGCCACAUAUACACCUCU	972
CCACCCCAAGAUUUAAA	235	CCACCCCAAGAUUUAAA	235	UUUAAUUCUUGUGGGGUGG	973
GUAAAAAUUGGAUGACAG	236	GUAAAAAUUGGAUGACAG	236	CUGUCAUCCAUUUUUUUAU	974
AUAUAGCAACAGACAUAC	237	AUAUAGCAACAGACAUAC	237	GUUUGUCUGUUGCUAUUAU	975
GCAUAUAGCAGCUGCUUU	238	GCAUAUAGCAGCUGCUUU	238	AAAGCAGCUGCUUAUAUGC	976
GGCAGGUGAUGAUUGUGUG	239	GGCAGGUGAUGAUUGUGUG	239	CACACAUAUACACCUGCC	977
AUGAUACAGUAUUAAGAA	240	AUGAUACAGUAUUAAGAA	240	UCUUCUAAUACUGUAUCAU	978
GAUGGCAGGUGAUGAUUGU	241	GAUGGCAGGUGAUGAUUGU	241	ACAAUCAUACCCUGCCAU	979
CAUAAUAGCAACAGACAU	242	CAUAAUAGCAACAGACAU	242	UAUGUCUGUUGCUAUUAUG	980
AAAAUUUCGGGUUUUAUUA	243	AAAAUUUCGGGUUUUAUUA	243	UAAUAAACCCGAAAAUUUU	981
ACAUAAUAGCAACAGACAU	244	ACAUAAUAGCAACAGACAU	244	AUGUCUGUUGCUAUUAUGU	982
AUUUCAAAAAUUGGGCCUG	245	AUUUCAAAAAUUGGGCCUG	245	CAGGCCAAUUUUUUGAAAU	983
CUGGAAAGGUGAAGGGGCA	246	CUGGAAAGGUGAAGGGGCA	246	UGCCCCUUCACCCUUUCCAG	984

AAAAAGGCGCAGGUGAU	247	AAAACAGAUGGCAGGGUGAU	247	AUACCUGGCCAUCUGUUUU	985
UUUCAAAAUUUGGCCUGA	248	UUUCAAAAUUUGGCCUGA	248	UCAGGCCCAAUUUUUGAAA	986
GAGAGAACCAAGGGGAAGU	249	GAGAGAACCAAGGGGAAGU	249	ACUCCCCCUUGGUUCUCUC	987
CUCUGGAAAGGUGAAGGGG	250	CUCUGGAAAGGUGAAGGGG	250	CCCUUACACCUUUCAGAG	988
AUUAGCAGGAAGAUUGCCA	251	AUUAGCAGGAAGAUUGCCA	251	UGGCCAUUCCUGCUAAU	989
GAGCCACCCCACAAGAUUU	252	GAGCCACCCCACAAGAUUU	252	AAAUUUGUGGGGUGGCUC	990
CAUAGCAGGAACUACUAGU	253	CAUAGCAGGAACUACUAGU	253	ACUAGUAGUUCUGCUAUG	991
UUUUUAAAAGAAAAGGGGG	254	UUUUUAAAAGAAAAGGGGG	254	CCCCCUUUUUUUUUUAAA	992
GCGGAGGCUAGAAGGAGAG	255	GCGGAGGCUAGAAGGAGAG	255	CUCUCCUUCUAGCCUCCGC	993
CAGUAUUAGUAGGACCUAC	256	CAGUAUUAGUAGGACCUAC	256	GUAGGUCCUACUAAUACUG	994
AGGGGGAUUUGGAGGUUUU	257	AGGGGGAUUUGGAGGUUUU	257	AAAACCCUCCAAUUCGCCU	995
ACAGUAUUAGUAGGACCUA	258	ACAGUAUUAGUAGGACCUA	258	UAGGUCCUACUAAUACUGU	996
GACUAGCGGAGGCUAGAAG	259	GACUAGCGGAGGCUAGAAG	259	CUUCUAGCCUCCGCUAGUC	997
GUUUUUACAGGGACAGCA	260	GUUUUUACAGGGACAGCA	260	UGCUCUCCUUGUAUAAAC	998
CAGGUGAUGAUUGUGGCG	261	CAGGUGAUGAUUGUGGCG	261	GCCACAUAUACUACCCUG	999
AGCGGAGGCUAGAAGGAGA	262	AGCGGAGGCUAGAAGGAGA	262	UCUCCUUCUAGCCUCCGCU	1000
UCUAUGUJAGUUGGGGCAGC	263	UCUAUGUJAGUUGGGGCAGC	263	GCUGCCCCAUUCUACAUGA	1001
UAAAAAUUGGAUGACAGA	264	UAAAAAUUGGAUGACAGA	264	UCUGUCAUCCAUUUUUUA	1002
GCAGCAGGAAGCACUAUGG	265	GCAGCAGGAAGCACUAUGG	265	CCAUAGUGCUUCCUGCUGC	1003
UUUUUACAGGGACAGCAGA	266	UUUUUACAGGGACAGCAGA	266	UCUGCUGUCCUGUAUUA	1004
AAACAGUUGCGAGGUGAUG	267	AAACAGUUGCGAGGUGAUG	267	CAUCACCCUGCCAUCUGUUU	1005
AUUCAAAAUUUUCGGGUUU	268	AUUCAAAAUUUUCGGGUUU	268	AAACCCGAAAAUUUUGAAU	1006
GGGGAUUGGAGGUUUUAU	269	GGGGAUUGGAGGUUUUAU	269	AUAAACCCUCCAAUUCCCC	1007
GCCACCCCACAAGAUUUA	270	GCCACCCCACAAGAUUUA	270	UUAAAUUUGUGGGGUGGC	1008
GAUGAUACAGUUAUJAGAG	271	GAUGAUACAGUUAUJAGAG	271	CUUCUAAUACUGUAUCAUC	1009
UAAUAGCAACAGACAUACA	272	UAAUAGCAACAGACAUACA	272	UGUAUGUCUGUUGCUAUUA	1010
GAGGCUAGAAGGAGAGAGA	273	GAGGCUAGAAGGAGAGAGA	273	UCUCUCUCCUUCUAGCCUC	1011
GUACAGUAUUJAGUAGGACC	274	GUACAGUAUUJAGUAGGACC	274	GGUCCUACUAAUACUGUAC	1012
UAGCGGAGGCUAGAAGGAG	275	UAGCGGAGGCUAGAAGGAG	275	CUCUUCUAGCCUCCGCUA	1013
CGGAGGCUAGAAGGAGAGA	276	CGGAGGCUAGAAGGAGAGA	276	UCUCUCCUUCUAGCCUCCG	1014
GGUACAGUAUUJAGUAGGAC	277	GGUACAGUAUUJAGUAGGAC	277	GUCCUACUAAUACUGUACC	1015
AAAUUUCGGGUUUUAUJAC	278	AAAUUUCGGGUUUUAUJAC	278	GUAAUAAACCCGAAAAUUU	1016
AGCAGCAGGAAGCACUUAUG	279	AGCAGCAGGAAGCACUUAUG	279	CAUAGUGCUUCCUGCUGCU	1017
AGCCACCCCACAAGAUUUA	280	AGCCACCCCACAAGAUUUA	280	UAAAUUUGUGGGGUGGCU	1018
AACCAAGGGGAAGUGACAU	281	AACCAAGGGGAAGUGACAU	281	AUGUCACUUCUCCUUGGUU	1019
AAGGGGAAGUGACAUAGCA	282	AAGGGGAAGUGACAUAGCA	282	UGCUAUGUCACUCCUCCU	1020

UUAAGCCAGGAUUGGAG	283	UUAAGCCAGGAUUGGAG	283	CAUCCAUUCCUGGCUUUA	1021
ACUAGCGGAGCUAGAAG	284	ACUAGCGGAGCUAGAAG	284	CCUUCUJAGCCUCCGCUAGU	1022
UAGGUACAGUUAUAGUAG	285	UAGGUACAGUUAUAGUAG	285	CCUACUAAUACUGUACCUA	1023
GGGGAAUUGGAGGUUUUA	286	GGGGAAUUGGAGGUUUUA	286	UAAAACCUCCAAUUCCCCC	1024
AGUUGGCAGGUGAUAUUG	287	AGUUGGCAGGUGAUAUUG	287	CAUACUACCCUGCCCAUCU	1025
UUAACAAUUGGCCAUUGAC	288	UUAACAAUUGGCCAUUGAC	288	GUCAUUGGCCAUUGUUUUA	1026
UGGCAGGUGAUAUUGUGU	289	UGGCAGGUGAUAUUGUGU	289	ACACAUAUACCCUGGCCA	1027
UAAAUAUAGCAGGAAGAU	290	UAAAUAUAGCAGGAAGAU	290	CAUCUUCUGCUAAUUA	1028
AGGAGCCACCCACAAAGAU	291	AGGAGCCACCCACAAAGAU	291	AUCUUGUGGGUGGCUCCU	1029
GUUUAGUAGGACCUACAC	292	GUUUAGUAGGACCUACAC	292	GUGUAGGUCCUACUAAUAC	1030
AAUCCCCAAAGUCAAGGAG	293	AAUCCCCAAAGUCAAGGAG	293	CUCCUUGACUUGGGGAUU	1031
CCAGGCCAGAUAGAGAGAC	294	CCAGGCCAGAUAGAGAGAC	294	GUUCUCUACUUGGCCUUG	1032
CCAUUGACAGGAAGAAAA	295	CCAUUGACAGGAAGAAAA	295	UUUUUUCUUCUGUCAUUGG	1033
CAGUUGGCAGGUGAUAUU	296	CAGUUGGCAGGUGAUAUU	296	AAUCAUACCCUGCCAUUCUG	1034
CAGAUAGAGAACCAAGGG	297	CAGAUAGAGAACCAAGGG	297	CCCUUGGUUCUCUCAUCUG	1035
GCCAUUGACAGGAAGAAAA	298	GCCAUUGACAGGAAGAAAA	298	UUUUUCUUCUGUCAUUGGC	1036
UAUUAGUAGGACCUACACC	299	UAUUAGUAGGACCUACACC	299	GGUGUAGGUCCUACUAAUA	1037
UCUCGACGAGACUCGCG	300	UCUCGACGAGACUCGCG	300	GCGAGUCCUGCGUCGAGA	1038
AGUAGAGAACCAAGGGG	301	AGUAGAGAACCAAGGGG	301	CCCUUGGUUCUCUCAUCU	1039
AUCCCCAAAGUCAAGGAGU	302	AUCCCCAAAGUCAAGGAGU	302	ACUCCUUGACUUGGGGAU	1040
AAUUAGCAGGAAGUUGGCC	303	AAUUAGCAGGAAGUUGGCC	303	GGCCAUUCUCCUGCUAAUU	1041
GGGAAUUGGAGGUUUUAUC	304	GGGAAUUGGAGGUUUUAUC	304	GAUAAAACCUCCAAUUC	1042
CUUGACGACGACUCGCGU	305	CUUGACGACGACUCGCGU	305	AGCCGAGUCCUGCGUCGAG	1043
AUGGCCAUUGACAGAAGAA	306	AUGGCCAUUGACAGAAGAA	306	UUCUUCUGUCAUUGGCCAU	1044
AAAAUUAGCAGGAAGUUG	307	AAAAUUAGCAGGAAGUUG	307	CCAUCUCCUGCUAAUUAU	1045
ACGACGACUCGCGUUGCU	308	ACGACGACUCGCGUUGCU	308	AGCAAGCCGAGUCCUGCGU	1046
UAAACAAUGGCCAUUGACA	309	UAAACAAUGGCCAUUGACA	309	UGUCAUUGGCCAUUGUUUA	1047
GAUGGAACAAGCCCCAGAA	310	GAUGGAACAAGCCCCAGAA	310	UUCUGGGCUUGUUCCAUC	1048
AAUGAACAGUAGUAAAU	311	AAUGAACAGUAGUAAAU	311	AUUUAUCUACUUGUUAU	1049
AUUUGAGGUUUUAUCAAAG	312	AUUUGAGGUUUUAUCAAAG	312	CUUUGAUAAACCUCCAAU	1050
AGGUAGAAAGGAGAGAGAU	313	AGGUAGAAAGGAGAGAGAU	313	AUCUCUCCUUCUJAGCCU	1051
AGAUUGGUUGCGAGCGUC	314	AGAUUGGUUGCGAGCGUC	314	GACGCUUCGACCCCAUCU	1052
AGGUACAGUUAUAGUAGGA	315	AGGUACAGUUAUAGUAGGA	315	UCCUACUAAUACUJAGCCU	1053
GGAGGCUAGAAGGAGAGAG	316	GGAGGCUAGAAGGAGAGAG	316	CUCUCUCCUUCUJAGCCUCC	1054
CAGGACAUAAACAAGGUAGG	317	CAGGACAUAAACAAGGUAGG	317	CCUACCUUGUUAUUGUCCUG	1055
AGUAUUAGUAGGACCUACA	318	AGUAUUAGUAGGACCUACA	318	UGUAGGUCCUACUAAUACU	1056

UUGACAGAAAGAAAAUAA	319	UUGACAGAAAGAAAAUAA	319	UUUUUUUUUUUUUUUUUU	1057
UGGAGAAUGAAUUUAUA	320	UGGAGAAUGAAUUUAUA	320	UUAUAAUUUACAUUCUCA	1058
CUCUCGACGCAGGACUGG	321	CUCUCGACGCAGGACUGG	321	CCGAGUCCUGCGUCGAGAG	1059
AUGAACAAAGUAGUAAAUU	322	AUGAACAAAGUAGUAAAUU	322	AAUUUAUCUACUUUUUCAU	1060
UGGCCAUUGACAGAAGAAA	323	UGGCCAUUGACAGAAGAAA	323	UUUUUUUUUUUUUUUUUU	1061
AUACCCAUUUUUCAGCAU	324	AUACCCAUUUUUCAGCAU	324	AUGCUGAAAAACAUGGGUUAU	1062
UUUAAAAGAAAAAGGGGGA	325	UUUAAAAGAAAAAGGGGGA	325	UCCCCUUUUUUUUUUUUAAA	1063
CGACGCAGGACUCGGCUUG	326	CGACGCAGGACUCGGCUUG	326	CAAGCCGAGUCCUGCGUCG	1064
AUUGACAGAAGAAAAAUAA	327	AUUGACAGAAGAAAAAUAA	327	UAUUUUUUUUUUUUUUUUUU	1065
CUAGAAAGGAGAGAUUGGG	328	CUAGAAAGGAGAGAUUGGG	328	CCCAUCUCUCUCCUUCUAG	1066
UGGCAGGAAGAAAGCGGAGA	329	UGGCAGGAAGAAAGCGGAGA	329	UCUCCGCUUCUUCUCCGCCA	1067
CAAUCCCCAAAGUCAAGGA	330	CAAUCCCCAAAGUCAAGGA	330	UCCUUGACUUUUGGGGAUUG	1068
AAAUUCAAUUUUUCGGGU	331	AAAUUCAAUUUUUCGGGU	331	ACCCGAAAAUUUUUGAAUUU	1069
GAUUUGGAGGUUUUAUCA	332	GAUUUGGAGGUUUUAUCA	332	UUGAUAAAAACCUCCAAUUC	1070
GACCGGAGUCGGCUUGC	333	GACCGGAGUCGGCUUGC	333	GCAAGCCGAGUCCUGCGUC	1071
UUUGACUAGCGGAGGCUAG	334	UUUGACUAGCGGAGGCUAG	334	CUAGCCUCCGCUAGUCAAA	1072
AUAGGUACAGUUAUUGUAG	335	AUAGGUACAGUUAUUGUAG	335	CUACUAAUACUAGUACCUAU	1073
GGCUJAGAAAGGAGAGAGUAG	336	GGCUJAGAAAGGAGAGAGUAG	336	CAUCUCUCUCCUUCUAGCC	1074
ACCAGGCCAGAUAGAGAAA	337	ACCAGGCCAGAUAGAGAAA	337	UUCUCUCAUCUGGCCUUGGU	1075
GAUGAGAGAAACCAAGGGGA	338	GAUGAGAGAAACCAAGGGGA	338	UCCCCUUGGUUUCUCUCAUC	1076
GGAGCAGCAGGAAGCACUA	339	GGAGCAGCAGGAAGCACUA	339	UAGUGCUUCCUGCUGCUCC	1077
UCUCUCGACGCAGGACUCG	340	UCUCUCGACGCAGGACUCG	340	CGAGUCCUGCGUCGAGAGA	1078
UCCCUACAUCUCCCCAAAGU	341	UCCCUACAUCUCCCCAAAGU	341	ACUUUUGGGGAUUGUAGGGA	1079
UUGGAGGUUUUAUCAAGU	342	UUGGAGGUUUUAUCAAGU	342	ACUUUGAUAAAAACCUCCAA	1080
ACUGUACCAAGUAAAAUUA	343	ACUGUACCAAGUAAAAUUA	343	UUAUUUUUACUUGGUACAGU	1081
AUGGCAGGUAGUUAUUGU	344	AUGGCAGGUAGUUAUUGU	344	CACAUAUCAACCUUGCCAU	1082
GAGGAAUUGAACAAAGUAGA	345	GAGGAAUUGAACAAAGUAGA	345	UCUACUUGUUAUUAUCCUC	1083
AGACAUAAUAGCAACAGAC	346	AGACAUAAUAGCAACAGAC	346	GUCUGUUGCUAUUAUUGUCU	1084
AAUUUAGCAGGAAGUUGC	347	AAUUUAGCAGGAAGUUGGC	347	GCCAUUUUCCUGCUAAUUU	1085
UUGGAGAAUGAAUUUAUAU	348	UUGGAGAAUGAAUUUAUAU	348	AUAUAAUUUACUUUCUCCAA	1086
UCGACGCAGGACUCGGCUU	349	UCGACGCAGGACUCGGCUU	349	AAGCCGAGUCCUUGCGUCGA	1087
AAAAUCAAUUUUUCGGG	350	AAAAUCAAUUUUUCGGG	350	CCCGAAAAUUUUGAAUUUU	1088
CAGGCCAAGUAGACAGACC	351	CAGGCCAAGUAGACAGACC	351	GGUUCUCUCAUCUGGCCUCG	1089
UACCCAUUUUUCAGCAU	352	UACCCAUUUUUCAGCAU	352	AAUGCUGAAAAACAUGGGUA	1090
ACACAUGCCUUGUACCCA	353	ACACAUGCCUUGUACCCA	353	UGGGUACACAGGCAUGUGU	1091
GGCCAUUGACAGAAGAAA	354	GGCCAUUGACAGAAGAAA	354	UUUUUUUUUUUUUUUUUU	1092



GAGCAGCAGGAAAGCACUAU	355	GAGCAGCAGGAAAGCACUAU	355	AUAGUGCUUCCUGCUGCUC	1093
CUGUACCCAGUAAAAUUAAA	356	CUGUACCCAGUAAAAUUAAA	356	UUUAAUUUUUACUGGUACAG	1094
GAAUUGAUGACAGCAUGUC	357	GAAUUGAUGACAGCAUGUC	357	GACAUGCUGUCAUAUUUC	1095
CAUUGACAGAGAAAAAU	358	CAUUGACAGAGAAAAAU	358	AUUUUUUUUUUCUGUCAUUG	1096
AAUUGAUGACAGCAUGUCA	359	AAUUGAUGACAGCAUGUCA	359	UGACAUGCUGUCAUAUUU	1097
GCUAGAGGAGAGAGAGAGG	360	GCUAGAGGAGAGAGAGAGG	360	CCAUCUCUCUCCUUCUAGC	1098
UAGGGAUUUAUGGAAAAACAG	361	UAGGGAUUUAUGGAAAAACAG	361	CUGUUUUUCCAUAAUCCCUA	1099
GAUUUUUAUGUAGAUUUCAG	362	GAUUUUUAUGUAGAUUUCAG	362	CUGAAAUUCUAUAUUUUC	1100
CUACACCUUGUCAACAUAAU	363	CUACACCUUGUCAACAUAAU	363	AUUUUGUUGACAGGGUGAG	1101
ACAGAUGGCAGGUGAUGAU	364	ACAGAUGGCAGGUGAUGAU	364	AUCAUCACCCUGCCAUUCUGU	1102
CCACAGGGAUUGGAAAGGAU	365	CCACAGGGAUUGGAAAGGAU	365	AUCCUUUCCAUUCCCUUGG	1103
UUAGGGAUUUAUGGAAAAACA	366	UUAGGGAUUUAUGGAAAAACA	366	UGUUUUUCCAUAAUCCCUAA	1104
AGAUGCUGCAUUAAGCAG	367	AGAUGCUGCAUUAAGCAG	367	CUGCUUAUAUGCAGCAUCU	1105
AAUAGCAACAGACAUACAA	368	AAUAGCAACAGACAUACAA	368	UUUGAUGUCUGUUUGCUAUU	1106
AAUUCAAAAUUUUCGGGUU	369	AAUUCAAAAUUUUCGGGUU	369	AACCCGAAAAUUUUGAAUU	1107
CAGACUCACAAUAUGCAUU	370	CAGACUCACAAUAUGCAUU	370	AAUGCAUUAUUGUGAGUCUG	1108
UAACCCUUGCAACAUAAU	371	UAACCCUUGCAACAUAAU	371	GAUUGAUUCCUAAUGCAUA	1109
UGGAGGAAUUGCAACAGUA	372	UGGAGGAAUUGCAACAGUA	372	AUUUUGUUGACAGGGUGUA	1110
ACCAAGGGAAUGAGACUA	373	ACCAAGGGAAUGAGACUA	373	UACUUGUUAUUUCCUCCA	1111
GAGAUUGGUGCGAGAGCGU	374	GAGAUUGGUGCGAGAGCGU	374	UAUGUCACUCCCUUUGGU	1112
UAUAGGUUACAGUUAUAGUA	375	UAUAGGUUACAGUUAUAGUA	375	ACGCUCUCGCCACCCUUCUC	1113
AUUAGGGAUUUAUGGAAAAAC	376	AUUAGGGAUUUAUGGAAAAAC	376	UACUAAUACUGUACCCUAA	1114
UGGCUUGGAAAGAUACCU	377	UGGCUUGGAAAGAUACCU	377	GUUUUCCAUAAUCCCUAAU	1115
GAGAGUUGGGUGCGAGAGC	378	GAGAGUUGGGUGCGAGAGC	378	AGGUACUUAUCCACAGCCA	1116
CCUACACCUUGUCAACAUAA	379	CCUACACCUUGUCAACAUAA	379	GCUCUCGCCACCCUUCUCUC	1117
CAGCAGUACAAUUGGCAGU	380	CAGCAGUACAAUUGGCAGU	380	UUUGUUGACAGGGUGUAGG	1118
GGCUGUGGAAAGAUACCUA	381	GGCUGUGGAAAGAUACCUA	381	ACUGCAUUAUUGUACUGCUG	1119
AGAAAUUAUGUAGAUUUUA	382	AGAAAUUAUGUAGAUUUUA	382	UAGGUACUUAUCCACAGCC	1120
GCCACCUUUUCCUAGUGUU	383	GCCACCUUUUCCUAGUGUU	383	UGAAAUUCUACUAAUUUUCU	1121
GAUGCUGCAUUAAGCAGC	384	GAUGCUGCAUUAAGCAGC	384	AACACUAGGCAAAAGGUGGC	1122
GCUAUAGGUUACAGUUAUAG	385	GCUAUAGGUUACAGUUAUAG	385	GCUGCUUAUUGCAGCAUC	1123
AACAGAUGGCAGGUGAUGA	386	AACAGAUGGCAGGUGAUGA	386	CUAAUACUGUACCUUAJAGC	1124
AUCACUCUUUGGCAACGAC	387	AUCACUCUUUGGCAACGAC	387	UCAUCACCUUGCCAUUCUGUU	1125
ACAUGCCUGUGUACCCACA	388	ACAUGCCUGUGUACCCACA	388	GUCGUUGCCAAAGAGUGAU	1126
ACAGCAGUACAAUUGGCAG	389	ACAGCAGUACAAUUGGCAG	389	UGUGGGUACACAGGGCAUGU	1127
	390		390	CUGCCAUUUUGUACUGCUGU	1128



AUGCAUUAGGAAUUAUUA	391	AUGCAUUAGGAAUUAUUA	391	UGAAUGAUUCCUAAUUGCAU	1129
AAUUGGAGGUUUUAUCAA	392	AAUUGGAGGUUUUAUCAA	392	UUUGAUAAAACCCUCAAUU	1130
UUGGAGGAAUUAACAAGU	393	UUGGAGGAAUUAACAAGU	393	ACUUGUUAUUUCCUCCAA	1131
AUUGGAGGAAUUAACAAG	394	AUUGGAGGAAUUAACAAG	394	CUUGUUAUUUCCUCCAAU	1132
AAAAUUAUUAUUUUCGG	395	AAAAUUAUUAUUUUCGG	395	CCGAAAAUUUUGAAUUUUU	1133
AGGUGAAGGGGCGAGUAG	396	AGGUGAAGGGGCGAGUAG	396	UACUACUGCCCCUUCACCU	1134
CUAUAGGUACAGUAUUAGU	397	CUAUAGGUACAGUAUUAGU	397	ACUAAUACUGUACCUAUAG	1135
AUUAAGGCCAGGAUUGGAU	398	AUUAAGGCCAGGAUUGGAU	398	AUCCAUUCCUGGCUUUAAU	1136
GGAGGAAUUAACAAGUAG	399	GGAGGAAUUAACAAGUAG	399	CUACUUGUUAUUUCCUCC	1137
AGCAGUACAAUUGGCAGUA	400	AGCAGUACAAUUGGCAGUA	400	UACUGCCAUUUUGUACUGCU	1138
AUCAGUACAAUUGUCUCC	401	AUCAGUACAAUUGUCUCC	401	GGAAGCACAUUUGUACUGAU	1139
UAUGGGGUACCUUGUGGA	402	UAUGGGGUACCUUGUGGA	402	UCCACACAGGUACCCCAUA	1140
AGAGAUUGGUGCGAGGCG	403	AGAGAUUGGUGCGAGGCG	403	CGCUCUCGCACCCCAUCUCU	1141
GGUGAAGGGGCGAGUAAU	404	GGUGAAGGGGCGAGUAAU	404	UUACUACUGCCCCUUCACC	1142
GUGAAGGGGCGAGUAAU	405	GUGAAGGGGCGAGUAAU	405	AUUACUACUGCCCCUUCAC	1143
CGCAGGACUGGCUUGCUG	406	CGCAGGACUGGCUUGCUG	406	CAGCAAGCCGAGUCCUGCG	1144
CACAUGCCUGUACCCAC	407	CACAUGCCUGUACCCAC	407	GUGGUACACACAGGCAUGUG	1145
GAGAGAGAUUGGUGCGAGA	408	GAGAGAGAUUGGUGCGAGA	408	UCUCGCACCCAUUCUCUC	1146
UAGAAAGGAGAGAUUGGU	409	UAGAAAGGAGAGAUUGGU	409	ACCCAUCUCUCUCCUUCUA	1147
CACAGGGAUGGAAAGGAUC	410	CACAGGGAUGGAAAGGAUC	410	GAUCCUUAUCCAUCCUUGUG	1148
GGCAGGAAGGCGGAGAC	411	GGCAGGAAGGCGGAGAC	411	GUCUCCGCUUCCUUCUCCGCC	1149
UCCCCAAAGUCAAGGAGUA	412	UCCCCAAAGUCAAGGAGUA	412	UACUCCUUAUACUUAUGGGGA	1150
CCUGUCAACAUAAUUGGAA	413	CCUGUCAACAUAAUUGGAA	413	UUCCAUUUAUUGUUGACAGG	1151
UAUCAGUACAAUUGUCUUC	414	UAUCAGUACAAUUGUCUUC	414	GAAGCACAUUUGUACUGAU	1152
UGAAGGGGCGAGUAAUA	415	UGAAGGGGCGAGUAAUA	415	UAUUAUACUGCCCCUUA	1153
CUCAGAUUCUCCAUUAAG	416	CUCAGAUUCUCCAUUAAG	416	CUUAUUAUGCAGCAUCUGAG	1154
ACAGGGAUGGAAAGGAUCA	417	ACAGGGAUGGAAAGGAUCA	417	UGAUCCUUAUCCAUCCUUGU	1155
AAGAAAAGGGGGAUUGGG	418	AAGAAAAGGGGGAUUGGG	418	CCCAAUCCCCCUUUUCUU	1156
UCAUUAAGGGGAUUAUGGAA	419	UCAUUAAGGGGAUUAUGGAA	419	UUUCCAUAAUCCCUAAUGA	1157
GAAGGAGAGAGAUUGGUGC	420	GAAGGAGAGAGAUUGGUGC	420	GCACCCAUUCUCUCCUUC	1158
GUUAAACAUAUUGGAAU	421	GUUAAACAUAUUGGAAU	421	UCAUUGGCCAUUUGUUAAC	1159
AUGGACAAGUAGACUGUAG	422	AUGGACAAGUAGACUGUAG	422	CUACAGUCUACUUGUCCAU	1160
UAGUAGAUUUCAGAGAAU	423	UAGUAGAUUUCAGAGAAU	423	AGUUCUCUGAAAUUCUACUA	1161
CUGUCAACAUAAUUGGAAG	424	CUGUCAACAUAAUUGGAAG	424	CUUCCAAUUUAUUGUAGACAG	1162
GGGGCAGUAGUAAUACAAG	425	GGGGCAGUAGUAAUACAAG	425	CUUGUAUUUAUACUUGCCCC	1163
CAUUAAGGGAUUAUGGAAA	426	CAUUAAGGGAUUAUGGAAA	426	UUUUCCAUAUACCCUAAUG	1164

(17000000)

GAACUACUAGUACCCUUA	427	GAACUACUAGUACCCUUA	427	UGAAGGGUACUAGUAGUUC	1165
GCAGGAAGCACUAUGGCG	428	GCAGGAAGCACUAUGGCG	428	CGCCCAUAGUGCUUCCUGC	1166
AAGGAGAGAGUAGGUGCG	429	AAGGAGAGAGUAGGUGCG	429	CGACCCCAUCUCUCUCCUU	1167
CAGGAUUGGAUGGCCCAAA	430	CAGGAUUGGAUGGCCCAAA	430	UUUGGGCAUCCAUUCCUG	1168
GGAAUUGAACAGUAGUA	431	GGAAUUGAACAGUAGUA	431	UAUCUACUUGUUAUUC	1169
AAAAGACACCAAGGAGCU	432	AAAAGACACCAAGGAGCU	432	AGCUUCCUUGGUGCUUUU	1170
AUCAUUAAGCACCAACCAG	433	AUCAUUAAGCACCAACCAG	433	CUGGUUGUGCUUGAAUGAU	1171
AACAAGUAGUAAUUAUAGU	434	AACAAGUAGUAAUUAUAGU	434	ACUAAUUUAUCUACUUGUU	1172
AGGAAUUGAACAGUAGAU	435	AGGAAUUGAACAGUAGAU	435	AUCUACUUGUUAUUCU	1173
GCAGGACUCGGCUUGCUGA	436	GCAGGACUCGGCUUGCUGA	436	UCAGCAAGCCGAGUCCUGC	1174
GAAUCAUUAAGCACCAACC	437	GAAUCAUUAAGCACCAACC	437	GGUUGUCUUGAAUGAUUC	1175
CCUCAGAUUCUGCAUAUAA	438	CCUCAGAUUCUGCAUAUAA	438	UUUAUUGCAGCAUCUGAGG	1176
GAUGGAAAGGAUACCCAGC	439	GAUGGAAAGGAUACCCAGC	439	GCUGGUGAUCCUUCUCCAU	1177
AGGAGAGAGAUUGGUGCGA	440	AGGAGAGAGAUUGGUGCGA	440	UCGCACCCCAUCUCUCUCCU	1178
CAUGGACAAGUAGACUGUA	441	CAUGGACAAGUAGACUGUA	441	UACAGUCUACUUGUCCAU	1179
UCAGAUUCUGCAUAUAAGC	442	UCAGAUUCUGCAUAUAAGC	442	GCUUUAUUGCAGCAUCUGA	1180
AUGGAGAAAUUAGUAGAU	443	AUGGAGAAAUUAGUAGAU	443	AUCUACUAAUUUUUCUCCAU	1181
GAGAAAUUAGUAGAUUUC	444	GAGAAAUUAGUAGAUUUC	444	GAAUUCUACUAAUUUUUCUC	1182
AUGCAGCAUGUCAGGGAG	445	AUGCAGCAUGUCAGGGAG	445	CUCCUGACAUUCUGUCU	1183
AGGCCAGUAGAGAACCA	446	AGGCCAGUAGAGAACCA	446	UGGUUCUCUACUUCUGGCCU	1184
AGAGAGAUUGGUGCGAG	447	AGAGAGAUUGGUGCGAG	447	CUCUCGACCCCAUCUCUCU	1185
ACCCAUGUUIUCAGCAUUA	448	ACCCAUGUUIUCAGCAUUA	448	UUAUGCUGAAAAACAUGGGU	1186
GAUGACAGCAUGUCAGGGA	449	GAUGACAGCAUGUCAGGGA	449	UCCUGACAUUCUGUCU	1187
AGCCAGGAUUGGAUGGCC	450	AGCCAGGAUUGGAUGGCC	450	GGCCCAUCCAUUCCUGGCU	1188
UGAUGACAGCAUGUCAGGG	451	UGAUGACAGCAUGUCAGGG	451	CCUUGACAUUCUGUCAUA	1189
CAGGAAGCACUAUGGCGC	452	CAGGAAGCACUAUGGCGC	452	GGCCCAUAGUGCUUCCUG	1190
ACAGACUCACAAUAUGCAU	453	ACAGACUCACAAUAUGCAU	453	AUGCAUAUUGUGAGUCUGU	1191
UGGAGGUUUUAUCAAAGUA	454	UGGAGGUUUUAUCAAAGUA	454	UACUUUGAUAAAACCCUCA	1192
AAGCCAGGAUUGGAUGGCC	455	AAGCCAGGAUUGGAUGGCC	455	GGCCAUCCAUUCCUGGCUU	1193
UUUUGACUAGCGGAGCUA	456	UUUUGACUAGCGGAGCUA	456	UAGCCUCCGCUAGUCAAAA	1194
CAGAUGCUGCAUAUAAGCA	457	CAGAUGCUGCAUAUAAGCA	457	UGCUAUAUUGCAGCAUCUG	1195
UUGGGCCUGAAAAUCCAU	458	UUGGGCCUGAAAAUCCAU	458	UAUGGAUUUUCAGGCCCAA	1196
GCAUGGACAAGUAGACUGU	459	GCAUGGACAAGUAGACUGU	459	ACAGUCUACUUGUCCAU	1197
ACCUUGUCAACAUAAUUGGA	460	ACCUUGUCAACAUAAUUGGA	460	UCCAAUUUUGUUGACAGGU	1198
CAGGAACUACUAGUACCCU	461	CAGGAACUACUAGUACCCU	461	AGGUACUAGUAGUCCUG	1199
AUAGCAACAGACAUACAAA	462	AUAGCAACAGACAUACAAA	462	UUUGUAGUGUCUUGUUAU	1200

GGAGAGAGAU	GGGUGCGAG	463	GGAGAGAGAU	GGGUGCGAG	463	CUCGACCCCAUCUCUCUC	1201
ACACCUCA	CAUAUUAU	464	ACACCUCA	CAUAUUAU	464	CAUAUUAU	1202
AGAAUUAU	GAGACGAGU	465	AGAAUUAU	GAGACGAGU	465	ACAUUGUC	1203
AGAAUUAU	GAGACGAGU	466	AGAAUUAU	GAGACGAGU	466	CACCAUCUCUCUCUCUC	1204
AAUUAUUA	CAAGCAACCA	467	AAUUAUUA	CAAGCAACCA	467	UGGUUGUCUUGAAUGAUU	1205
CAAAUUAU	GGGCGGAA	468	CAAAUUAU	GGGCGGAA	468	UUUUCAGGCCCAUUAUUG	1206
GCAGUACA	UAGGAGUUA	469	GCAGUACA	UAGGAGUUA	469	AUACUGCCAUUUGUACUGC	1207
GGGAGUA	UAUACAGA	470	GGGAGUA	UAUACAGA	470	UCUUGUAUUAUACUGGCC	1208
UCAUUAAG	CAACACAGA	471	UCAUUAAG	CAACACAGA	471	UCUGGUUGUCUUGAAUGA	1209
AUGAUGAG	CAUUGCAGG	472	AUGAUGAG	CAUUGCAGG	472	CCUGACAUGCUGUCAU	1210
GAACAAGU	AGAUAAUUAU	473	GAACAAGU	AGAUAAUUAU	473	CUAAUUUAUCUACUUGUUC	1211
UGACAGCAU	GUCAGGAGU	474	UGACAGCAU	GUCAGGAGU	474	ACUCCUGACAUGCUGUCA	1212
GGAACUAU	UAGCCUUC	475	GGAACUAU	UAGCCUUC	475	GAGGGUACUAGUAGUUC	1213
CACCUCAU	CAUAUUAU	476	CACCUCAU	CAUAUUAU	476	CCAAUUAUUGUACAGGUG	1214
GGCCAGAU	GAGAGAACCA	477	GGCCAGAU	GAGAGAACCA	477	UUGGUUCUCUCUCUGGCC	1215
UGUGUAC	CACAGACCCCA	478	UGUGUAC	CACAGACCCCA	478	UGGGGUCUGUGGGUACACA	1216
GGAUUAU	CAAGCACAAC	479	GGAUUAU	CAAGCACAAC	479	GUUGUGCUUGAAUGAUUCC	1217
CAGUACA	UUGGAGUUAU	480	CAGUACA	UUGGAGUUAU	480	AAUACUGCCAUUUGUACUG	1218
GCAGGAAG	AGCGGAGACA	481	GCAGGAAG	AGCGGAGACA	481	UGUCUCCGCUUCCUUGC	1219
AAAGCCAG	GAUUGGAGC	482	AAAGCCAG	GAUUGGAGC	482	GCCAUCUCCUUGGCGUUA	1220
UGAACAAU	GAUAAUUA	483	UGAACAAU	GAUAAUUA	483	UAAUUUAUCUACUUGUUA	1221
CAAAAUUA	CAAAUUAU	484	CAAAAUUA	CAAAUUAU	484	CGAAAAUUUGAAUUUUUG	1222
UAGGACCU	UACCCUGUCA	485	UAGGACCU	UACCCUGUCA	485	UUGACAGGUGUAGGUCCUA	1223
GCCAGAU	GAGAGAACCAAG	486	GCCAGAU	GAGAGAACCAAG	486	CUUGGUUCUCUACUUGGC	1224
GACAGCU	GGACUUGCAU	487	GACAGCU	GGACUUGCAU	487	CAUUGACAGUCCAGCUGUC	1225
AAAGCCAC	CUUUGCCUAGU	488	AAAGCCAC	CUUUGCCUAGU	488	ACUAGGCAAGGUGGCUUU	1226
GAAUUAU	CAAGUAGUUA	489	GAAUUAU	CAAGUAGUUA	489	UUAUCUACUUGUUAUUAU	1227
ACAAUUUA	CAAGAAAGG	490	ACAAUUUA	CAAGAAAGG	490	CCUUUUCUUUUAUUUUUG	1228
GCUGUGGA	GAUACCUAA	491	GCUGUGGA	GAUACCUAA	491	UUAGGUUAUUCUUCCACAGC	1229
UGUCAACA	UAUUGGAAGA	492	UGUCAACA	UAUUGGAAGA	492	UCUUCCAUAUUGUUGACA	1230
UAAAGAAA	GGGGGAU	493	UAAAGAAA	GGGGGAU	493	AAUCCCCCUUUUUUUUA	1231
CAUUUUUA	AAAGAAAGG	494	CAUUUUUA	AAAGAAAGG	494	CCUUUUUUUUUUUUUUUUG	1232
UUAGUAGU	UAUUCAGAGAC	495	UUAGUAGU	UAUUCAGAGAC	495	GUUCUCUGAAAUUCUACUAA	1233
AAUUUUUA	AAAGAAAGG	496	AAUUUUUA	AAAGAAAGG	496	CCCCUUUUUUUUUUUUU	1234
UAGCAAC	AGACAUACAAC	497	UAGCAAC	AGACAUACAAC	497	GUUUGUAUGUCUGUUGCUA	1235
UGGAACA	AGCCCGAGAAGA	498	UGGAACA	AGCCCGAGAAGA	498	UCUUCUGGGGCUUGUUCUA	1236

(400/099)

AGGAUGAGGAAUUAAGAACA	499	AGGAUGAGGAAUUAAGAACA	499	AUGUUCUAAUCCUAUCCU	1237
GACAAUUGGAGAGUGAAU	500	GACAAUUGGAGAGUGAAU	500	AUUCACUUCUCCAAUUGUC	1238
ACAGACCCCAACCCACAAG	501	ACAGACCCCAACCCACAAG	501	CUUGUGGGUUGGGGUCUGU	1239
CACCUAGAACUUUAAUUGC	502	CACCUAGAACUUUAAUUGC	502	GCAUUUAAAGUUCUAGGUG	1240
GAGCAACACCCCAACAG	503	GAGCAACACCCCAACAG	503	CUGGUGGGGUCUUGGCUC	1241
AGGACCUACACCCGUCAC	504	AGGACCUACACCCGUCAC	504	GUUGACAGGUGUAGGUCCU	1242
UUACAAAAUUCAAAAUUU	505	UUACAAAAUUCAAAAUUU	505	AAAUUUUGAAUUUUUGUAA	1243
GGAGGUUUUAUCAAAGUAA	506	GGAGGUUUUAUCAAAGUAA	506	UUACUUUGAAUAAACCUCC	1244
CUGGCUUGGAAAGAUACC	507	CUGGCUUGGAAAGAUACC	507	GGUAUCUUUCCACAGCCAG	1245
GGAGAUGUAAUUAUAA	508	GGAGAUGUAAUUAUAA	508	UUAAUAAUUCACUUCUC	1246
AUGAUGACAGCAUGUCAG	509	AUGAUGACAGCAUGUCAG	509	CUGACAUUGCUGUCAUUAU	1247
AUCAUUAGGGAAUUAUGGAA	510	AUCAUUAGGGAAUUAUGGAA	510	UUCCAUAAUCCCUAAUGAU	1248
UCAAUUUUGGGCCUGAAA	511	UCAAUUUUGGGCCUGAAA	511	UUUCAGGGCCCAUUUUUGA	1249
ACCUACACCGUCAAUA	512	ACCUACACCGUCAAUA	512	UAUGUUAGACAGGUGUAGGU	1250
GAUGAGGAUUAGAAUUGG	513	GAUGAGGAUUAGAAUUGG	513	CCAUGUUUUAUCCUAUC	1251
ACAGCUGGACUGUCAAUGA	514	ACAGCUGGACUGUCAAUGA	514	UCAUUGACAGUCCAGCUGU	1252
CCUCAGAUUGCUCAAUA	515	CCUCAGAUUGCUCAAUA	515	UAUAUGCAGCAUCUGAGGG	1253
AUUAGUAUUUCAGAGAA	516	AUUAGUAUUUCAGAGAA	516	UUCUCUGAAAUUCUACUAAU	1254
AGAAAGAGCAGAGACAGU	517	AGAAAGAGCAGAGACAGU	517	ACUGUCUUCUGCUCUUCU	1255
GACCUACACCCUGUCAAU	518	GACCUACACCCUGUCAAU	518	AUGUUGACAGGUGUAGGUC	1256
CACUCUUUGCAACGACCC	519	CACUCUUUGCAACGACCC	519	GGGUCGUUGCCAAAGAGUG	1257
AUGAGGAUUAGAAUUGGA	520	AUGAGGAUUAGAAUUGGA	520	UCCAUGUUCUAUCCUCAU	1258
AUUUUAAAAGAAAAGGGGG	521	AUUUUAAAAGAAAAGGGGG	521	CCCCUUUUUCUUUUAAAAU	1259
AGAACUUUAAAUGCAUGGG	522	AGAACUUUAAAUGCAUGGG	522	CCAUGCAUUUUAAGUUCU	1260
AUCUAUCAAUACAUGGAUG	523	AUCUAUCAAUACAUGGAUG	523	CAUCCAUGUAUUUGAUAGAU	1261
AUGGAACAAGCCCAAGAAG	524	AUGGAACAAGCCCAAGAAG	524	CUUCUGGGGCUUGUUCCAU	1262
UAUGACCCCAUCAAAGAC	525	UAUGACCCCAUCAAAGAC	525	GUCUUUUUGAUGGGUCAUAA	1263
CACAAUUUAAAAGAAAAG	526	CACAAUUUAAAAGAAAAG	526	CUUUUCUUUUAAAAUUGUG	1264
GAACUUUAAUUGCAUGGGU	527	GAACUUUAAUUGCAUGGGU	527	ACCAUGCAUUUAAAAGUUC	1265
AAAAGAAAAGGGGGAUUG	528	AAAAGAAAAGGGGGAUUG	528	CAUCCCCCUUUUCUUUU	1266
GGAUGGAAAGGAUACACAG	529	GGAUGGAAAGGAUACACAG	529	CUGGUGAUCCUUUCCAUCC	1267
AGGGGCAGUAGUAAUACAA	530	AGGGGCAGUAGUAAUACAA	530	UUGUAUUACUACUGCCCCU	1268
AAAGGGGGAUUGGGGGGU	531	AAAGGGGGAUUGGGGGGU	531	ACCCCCAAUCCCCCUUU	1269
AAGGGGGGAUUGGGGGGUA	532	AAGGGGGGAUUGGGGGGUA	532	UACCCCCAAUCCCCCUUU	1270
CAGGAUGAGGAUUAAGAACA	533	CAGGAUGAGGAUUAAGAACA	533	UGUUCUAAUCCUCAUCCUG	1271
AAAAUUAGUAUUUUCAGA	534	AAAAUUAGUAUUUUCAGA	534	UCUGAAAUCUACUAAUUUU	1272

(400/099)

GAUUUGGAGGAAUUGAAC	535	GAUUUGGAGGAAUUGAAC	535	UGUUCAUUUCCUCCAAUUC	1273
UACAAAAUUCAAAAUUUU	536	UACAAAAUUCAAAAUUUU	536	AAAAUUUGAAUUUUUGUA	1274
AGGAACUACUAGUACCCUU	537	AGGAACUACUAGUACCCUU	537	AAGGUACUAGUAGUCCU	1275
AAAGAAAAGGGGGAUUGG	538	AAAGAAAAGGGGGAUUGG	538	CCAAUCCCCCUUUUCUUU	1276
AAAAUUGGAUGACAGAAA	539	AAAAUUGGAUGACAGAAA	539	UUUCUGUACUCCAAUUUUU	1277
ACAGGAUGAGGAUUAGAAC	540	ACAGGAUGAGGAUUAGAAC	540	GUUCUAAUCCUACUCCUGU	1278
ACAAUUGGAGGAGUAAUU	541	ACAAUUGGAGGAGUAAUU	541	AAUUCACUUCUCCAAUUGU	1279
GGAUGAGGAUUAGAACAUG	542	GGAUGAGGAUUAGAACAUG	542	CAUGUUCUAAUCCUCAUCC	1280
UCACCUAGAACUUUAAAUG	543	UCACCUAGAACUUUAAAUG	543	CAUUUAAAGUUUCUAGGUGA	1281
AUUGGGCCUGAAAAUCCAU	544	AUUGGGCCUGAAAAUCCAU	544	AUGGAUUUUCAGGCCCAAU	1282
AAUUGGGCCUGAAAAUCCA	545	AAUUGGGCCUGAAAAUCCA	545	UGGAUUUUCAGGCCCAAUU	1283
GGACCUACACCCUGUCAACA	546	GGACCUACACCCUGUCAACA	546	UGUUGACAGGUGUAGGUCC	1284
GACAGGAUGAGGAUUAGAA	547	GACAGGAUGAGGAUUAGAA	547	UUCUAAUCCUCAUCCUGUC	1285
UCUAUCAUACAUGGAUGA	548	UCUAUCAUACAUGGAUGA	548	UCAUCCAUGAUUUUGAUAGA	1286
GGAUUUGGAGGAAUAGAAC	549	GGAUUUGGAGGAAUAGAAC	549	GUUCAUUUCCUCCAAUUC	1287
AAAAGGGGGAUUGGGGGG	550	AAAAGGGGGAUUGGGGGG	550	CCCCCAAUCCCCCUCUUU	1288
AAAAUUGGAUGACAGAAAC	551	AAAAUUGGAUGACAGAAAC	551	GUUUCUGUCAUCCAAUUUU	1289
CAAUUGGAGAGUAAUUA	552	CAAUUGGAGAGUAAUUA	552	UAAUUCACUUCUCCAAUUG	1290
AUGACCCAUCAAAAAGACUU	553	AUGACCCAUCAAAAAGACUU	553	AAGUCUUUUUGAUGGGUCAU	1291
CUUAAGCCUCAUAAAGCU	554	CUUAAGCCUCAUAAAGCU	554	AGCUUUUUUGAGGCUUAAAG	1292
AGUACA AUGUCUCCACA	555	AGUACA AUGUCUCCACA	555	UGUGAAGCACAUUUGUACU	1293
UUUCCGUGGGGACUUUCC	556	UUUCCGUGGGGACUUUCC	556	GGAAAGUCCCCAGCGGAAA	1294
CAGACAUACAACUAAAGA	557	CAGACAUACAACUAAAGA	557	UCUUUAGUUUUGUAUGUCUG	1295
UUAAGCCUCAUAAAGCUU	558	UUAAGCCUCAUAAAGCUU	558	AAGCUUUUUUGAGGCUUAA	1296
GGACAAUUGGAGAGUAGAA	559	GGACAAUUGGAGAGUAGAA	559	UUCACUUCUCCAAUUGUCC	1297
GGAUUGGGGGUACAGUGC	560	GGAUUGGGGGUACAGUGC	560	GCACUGUACCCCCCAAUUC	1298
AAAUUGGGCCUGAAAAUCC	561	AAAUUGGGCCUGAAAAUCC	561	GGAUUUUCAGGGCCCAAUUU	1299
GGGGAUUUGGGGGUACAG	562	GGGGAUUUGGGGGUACAG	562	CUGUACCCCCCAAUCCCCC	1300
GUUGGGGGACAUCAAGCAG	563	GUUGGGGGACAUCAAGCAG	563	CUGCUUGAUGUCCCCCCAC	1301
UCCUGGUGUGGAAAGAU	564	UCCUGGUGUGGAAAGAU	564	UAUCUUUCCACAGCCAGGA	1302
ACAAAAUUCAAAAUUUUC	565	ACAAAAUUCAAAAUUUUC	565	GAAAAUUUUGAAUUUUUGU	1303
GGGAUUGGGGGUACAGU	566	GGGAUUGGGGGUACAGU	566	ACUGUACCCCCCAAUCCCC	1304
UAAACACAGUGGGGGACA	567	UAAACACAGUGGGGGACA	567	UGUCCCCCACUGUGUUUA	1305
CAGACCCCAACCCACAAGA	568	CAGACCCCAACCCACAAGA	568	UCUUGUGGGUUGGGGUCUG	1306
AGGGCAAUUGGUACAUCA	569	AGGGCAAUUGGUACAUCA	569	UGAUGUACCAUUUGCCCCU	1307
AAUUGGAGGAAUUGAACAA	570	AAUUGGAGGAAUUGAACAA	570	UUGUUCAUUCCUCCAAUU	1308

(400/099)

AAGCCACCUUUGCCUAGUG	571	AAGCCACCUUUGCCUAGUG	571	CACUAGGCAAAAGGUGGCUU	1309
CCAUGUUUUCAGCAUUUUC	572	CCAUGUUUUCAGCAUUUUC	572	GAUAAUGCUGAAACACUAGG	1310
AAAGAAAAAUUCAGUAACA	573	AAAGAAAAAUUCAGUAACA	573	UGUUCUGAUUUUUUUUUUU	1311
AAAAAUUGGAGUCAGAA	574	AAAAAUUGGAGUCAGAA	574	UUCUGUCAUCCAAUUIUUUU	1312
CAGUACAAUUGCUUCCAC	575	CAGUACAAUUGCUUCCAC	575	GUGGAAGCACAUUGUACUG	1313
CUUCCGCGUGGGACUUC	576	CUUCCGCGUGGGACUUC	576	GAAAGUCCCCAGCGGAAAG	1314
GCAACGACAUACAAACUA	577	GCAACGACAUACAAACUA	577	UAGUUUGUAUGUCUGUUGC	1315
UAUCACCUAGAACUUAUA	578	UAUCACCUAGAACUUAUA	578	UUUAAAGUUUCUAGGUGAUA	1316
ACCCACAGACCCCAACCCA	579	ACCCACAGACCCCAACCCA	579	UGGGUUGGGGUCUGUGGGU	1317
GAUAGAUUGGAACAAGCCCC	580	GAUAGAUUGGAACAAGCCCC	580	GGGCUUUGUCCAUUCUAUC	1318
GCUUAAAGCCUCAAUAAAGC	581	GCUUAAAGCCUCAAUAAAGC	581	GCUUUAUUGAGGCUUUAAGC	1319
AUUGGGGGUACAGUGCAG	582	AUUGGGGGUACAGUGCAG	582	CUGACUGUACCCCCCAAU	1320
CCCACAGACCCCAACCCAC	583	CCCACAGACCCCAACCCAC	583	GUGGUUUGGGUUCUGUGGG	1321
AAAAUUGGGCCUGAAAAUC	584	AAAAUUGGGCCUGAAAAUC	584	GAUUUUCAGGCCCCAAUUUU	1322
CAUUCACGACACACCAGAU	585	CAUUCACGACACACCAGAU	585	AUCUGGUUGUGCUUGAAUG	1323
ACUUUAAUUGCAUGGUAA	586	ACUUUAAUUGCAUGGUAA	586	UUACCAUGCAUUUUAAGU	1324
UAGAACUUUAAUUGCAUGG	587	UAGAACUUUAAUUGCAUGG	587	CCAUGCAUUUAAAGUUCUA	1325
CUUUAAUUGCAUGGGUAAA	588	CUUUAAUUGCAUGGGUAAA	588	UUUACCCAUUGCAUUUUAAG	1326
GGGAUUGGGGGUACAGUG	589	GGGAUUGGGGGUACAGUG	589	CACUGUACCCCCCAAUCCC	1327
UAUGACCCCAUCAAAGACU	590	UAUGACCCCAUCAAAGACU	590	AGUCUUUUGAUGGGUCAA	1328
GAAGAAGCGGAGACAGCGA	591	GAAGAAGCGGAGACAGCGA	591	UCGCGUGUCUCCGCUUCUUC	1329
CCCAUGUUUUCAGCAUUAU	592	CCCAUGUUUUCAGCAUUAU	592	AUAAUGCUGAAACACUAGG	1330
AGGAUUGGAGGAAUUGAA	593	AGGAUUGGAGGAAUUGAA	593	UUCAUUUUCCUCCAAUUCU	1331
AGAGACAGGCUAAUUGUCAG	594	AGAGACAGGCUAAUUGUCAG	594	AAAAAUUAGCCUGUCUCU	1332
AAGUUAUCAGCAUUUAUCAG	595	AAGUUAUCAGCAUUUAUCAG	595	CUGACUAAUUUAUUCUACUU	1333
UUAUUGUCUGGUAUAGUGC	596	UUAUUGUCUGGUAUAGUGC	596	CUGAUAAUGCUGAAACAU	1334
AUUACAAAAUUCAAAAUU	598	AUUACAAAAUUCAAAAUU	598	AAUUUGAAUUUUUGUAAU	1336
GCCAGGAUUGGAGGCCCA	599	GCCAGGAUUGGAGGCCCA	599	UGGGCAUCCAUUCCUGGC	1337
CCUGGCUGUGGAAAGAUAC	600	CCUGGCUGUGGAAAGAUAC	600	GUUUCUUUCCACAGCCAGG	1338
UGUUUUCAGCAUUUAUCAGA	601	UGUUUUCAGCAUUUAUCAGA	601	UCUGAUAAUGCUGAAACA	1339
ACCUAGAACUUUAAUUGCA	602	ACCUAGAACUUUAAUUGCA	602	UGCAUUUAAAGUUUCUAGGU	1340
GGGAUGGAAAGGAUACACCA	603	GGGAUGGAAAGGAUACACCA	603	UGGUGAUCCUUUCCAUCCC	1341
AAUUAAAGCCAGGAUUGGA	604	AAUUAAAGCCAGGAUUGGA	604	UCCAUUCCUGGCUUUUAUU	1342
AAAGGAUUUGGAGGAAUUG	605	AAAGGAUUUGGAGGAAUUG	605	CAUUUCCUCCAAUUCUUUU	1343
ACUUUCCGCGUGGGACUUU	606	ACUUUCCGCGUGGGACUUU	606	AAAGUCCCGAGCGGAAAGU	1344

(400/099)

ACAGAAAGAAAAAUAAAAAG	607	ACAGAAGAAAAAUAAAAAG	607	CUUUUAUUUUUUUCUUCUGU	1345
AGCAACAGACAUACAAACU	608	AGCAACAGACAUACAAACU	608	AGUUUGUAUGUCUGUUGCU	1346
UAUUGUCUGUAUAGUGCA	609	UAUUGUCUGUAUAGUGCA	609	UGCACUAUACCGACAAUA	1347
UUAAAAGAAAGGGGGAU	610	UUAAAAGAAAGGGGGAU	610	AUCCCCCUUUUCUUUUA	1348
UGCUAAGCCUCAUAAAG	611	UGCUAAGCCUCAUAAAG	611	CUUUUAUGAGGCUUUAAGCA	1349
CAGGAAGUGGCCAGUAAA	612	CAGGAAGUGGCCAGUAAA	612	UUUACUGGCCAUUCUCCUG	1350
CCAGAUAGAGAACCAAGG	613	CCAGAUAGAGAACCAAGG	613	CCUUGGUUCUCUCAUGGG	1351
GAUUGGGGGUACAGUGCA	614	GAUUGGGGGUACAGUGCA	614	UGCACUGUACCCCCCAUUC	1352
AAUUGAACAAAGUAGUAAA	615	AAUUGAACAAAGUAGUAAA	615	UUUAUCUAUUGUUCAUUU	1353
AGCCACCUUUGCCUAGUGU	616	AGCCACCUUUGCCUAGUGU	616	ACACUAGGCAAGGUGGCU	1354
GACUUUCCGUGGGGACUU	617	GACUUUCCGUGGGGACUU	617	AAGUCCCCAGCGGAAAGUC	1355
CCAGUAAAUAUAAAGCCAG	618	CCAGUAAAUAUAAAGCCAG	618	CUGGCUUUAAUUUUACUGG	1356
GCAUUGUAUGCCCCUCCCA	619	GCAUUGUAUGCCCCUCCCA	619	UGGAGGGGCAUACAUUGC	1357
AACUUUAAUUGCAUGGGUA	620	AACUUUAAUUGCAUGGGUA	620	UACCCAUGCAUUUAAAGUU	1358
UUGGGGGUACAGUGCAGG	621	UUGGGGGUACAGUGCAGG	621	CCUGCACUGUACCCCCCAA	1359
GGACUUCCGUGGGGACU	622	GGACUUCCGUGGGGACU	622	AGUCCCCAGCGGAAAGUCC	1360
CUAGAACUUUAAUUGCAUG	623	CUAGAACUUUAAUUGCAUG	623	CAUGCAUUUAAAGUUCUAG	1361
UCAGUACAUGUGCUUCCA	624	UCAGUACAUGUGCUUCCA	624	UGGAAGCACAUUGUACUGA	1362
AAGGAUUGGAGGAAUAGA	625	AAGGAUUGGAGGAAUAGA	625	UCAUUUCCUCCAUUUCCUU	1363
UACCCACAGACCCCAACCC	626	UACCCACAGACCCCAACCC	626	GGGUUGGGGUCUGUGGGUA	1364
GAGACAGGCUAAUUUUUA	627	GAGACAGGCUAAUUUUUA	627	UAAAAAUUAGCCUGUCUC	1365
CUGCUUAGGCCUCAUAAA	628	CUGCUUAGGCCUCAUAAA	628	UUUAUUAGAGGCUUAAAGCAG	1366
AGGAAGAUUGGCCAGUAAAA	629	AGGAAGAUUGGCCAGUAAAA	629	UUUUACUGGCCAUUCUCCU	1367
AGACAUACAAACUAAAGAA	630	AGACAUACAAACUAAAGAA	630	UUCUUUAGUUUGUAGUCU	1368
CAUGUUUUCAGCAUUAUCA	631	CAUGUUUUCAGCAUUAUCA	631	UGAUAAUGCUGAATAACAU	1369
UUGGAAAGGACAGCAAAG	632	UUGGAAAGGACAGCAAAG	632	CUUUGCUGGUCUUCUCCAA	1370
GGCUGUUGGAAUUGUGGAA	633	GGCUGUUGGAAUUGUGGAA	633	UUCACAUUUCCAACAGCC	1371
UAAUUGGAGAAAUUAGUA	634	UAAUUGGAGAAAUUAGUA	634	UACUAAUUUUCUCCAUUA	1372
AGGAAGAGCGGAGACAGC	635	AGGAAGAGCGGAGACAGC	635	GCUGUCUCCGCUUCUCCU	1373
AAAAAGAAAAAUACAGUA	636	AAAAAGAAAAAUACAGUA	636	UACUGAUUUUUUCUUUUUU	1374
AUCAGAAAGAACCUCCAUU	637	AUCAGAAAGAACCUCCAUU	637	AAUGGAGGUUCUUCUGAU	1375
AGACCCCAACCCACAAGAA	638	AGACCCCAACCCACAAGAA	638	UUCUUUGGGUUGGGGUCU	1376
CAAGUAGAUAAUUUAGUCA	639	CAAGUAGAUAAUUUAGUCA	639	UGACUAAUUUAUCUACUUG	1377
AAAGCUAUAGGUACAGUUA	640	AAAGCUAUAGGUACAGUUA	640	AUACUGUACCUUAGCUUU	1378
UGCUGCAUUAAGCAGCUG	641	UGCUGCAUUAAGCAGCUG	641	CAGCUGCUUUAUUGCAGCA	1379
UUUAAUUGCAUGGGUAAAA	642	UUUAAUUGCAUGGGUAAAA	642	UUUUACCCCAUGCAUUUAAA	1380



(400/099)

UUUUCAGCAUUUUAUCAGAAG	643	UUUUCAGCAUUUUAUCAGAAG	643	CUUCUGAUAAUUCUGAAAA	1381
ACUGCUUUAAGCCUCAAUAA	644	ACUGCUUUAAGCCUCAAUAA	644	UUUUUGAGGCUUUAAGCAGU	1382
GGAAGGACCAAGCAAAAGCU	645	GGAAGGACCAAGCAAAAGCU	645	AGCUUUGCUGGCUUUUCC	1383
UGUACCAAGUAAAUUAAAG	646	UGUACCAAGUAAAUUAAAG	646	CUUUUUUUUUAUCUGGUACA	1384
GGAAGAAAAUUAAAAGCAU	647	GGAAGAAAAUUAAAAGCAU	647	AUGUUUUUUUUUUUUUUUC	1385
GUGUACCCACAGACCCCAA	648	GUGUACCCACAGACCCCAA	648	UUGGGGUCUGUGGUUACAC	1386
GGGGGAUUGGGGGUACA	649	GGGGGAUUGGGGGUACA	649	UGUACCCCAAUUCCCCCC	1387
GGAAGAAGCGGAGACGCG	650	GGAAGAAGCGGAGACGCG	650	CGCUGUCUCCGCUUUUCC	1388
GAAGCGGAGACAGCGACGA	651	GAAGCGGAGACAGCGACGA	651	UCGUCGCUUCCGCUUC	1389
UUAAUUGCAUGGGUAAAAG	652	UUAAUUGCAUGGGUAAAAG	652	CUUUUACCCAUUGCAUUUAA	1390
AACCCACUGCUUAAAGCCUC	653	AACCCACUGCUUAAAGCCUC	653	GAGGCUUUAAGCAGUGGGUU	1391
GUUUUCAGCAUUUUAUCAGAA	654	GUUUUCAGCAUUUUAUCAGAA	654	UUUCUGAUAAUUCUGAAAAC	1392
GGAUUUAAAUAUUUAAGUAA	655	GGAUUUAAAUAUUUAAGUAA	655	UUACUAAUUUUUUUUUAAUCC	1393
GUACCCACAGACCCCAACC	656	GUACCCACAGACCCCAACC	656	GGUUGGGGUCUGUGGGUAC	1394
GAUUAAAUAAAUAAGUAA	657	GAUUAAAUAAAUAAGUAA	657	CUUACUAAUUUUUUUUUAAUCC	1395
AAGCCUCAUAAAAGCUUUC	658	AAGCCUCAUAAAAGCUUUC	658	GCAAGCUUUUAUUGAGGCUU	1396
GCAGGACAUAAACAGGUAG	659	GCAGGACAUAAACAGGUAG	659	CUACCUUUAUUGUCCUGC	1397
CCCACUGCUUAAAGCCUCAA	660	CCCACUGCUUAAAGCCUCAA	660	UUGAGGCUUUAAGCAGUGGG	1398
GGGACUUCGCGUGGGGAC	661	GGGACUUCGCGUGGGGAC	661	GUCCCCAGCGGAAAGUCCC	1399
AUCACCUAGAACUUAUAAU	662	AUCACCUAGAACUUAUAAU	662	AUUUAAAAGUUCUAGGUGAU	1400
UAGAGCCUUGGAAGCAUCC	663	UAGAGCCUUGGAAGCAUCC	663	GGAUGCUUCCAGGGCUCUA	1401
GGGCUUUGGAAUUGUGGA	664	GGGCUUUGGAAUUGUGGA	664	UCCACAUUUCCAACAGCCC	1402
UUUCAGCAUUUUAUCAGAAG	665	UUUCAGCAUUUUAUCAGAAG	665	CCUUCUGAUAAUUGCUGAAA	1403
UGACCCAUCAAAAAGACUUA	666	UGACCCAUCAAAAAGACUUA	666	UAAGUCUUUUUUGAUGGGUCA	1404
AGAAAAAUAAAAGCAUUA	667	AGAAAAAUAAAAGCAUUA	667	UAAUGCUUUUUUUUUUUUCU	1405
AGAAGCGGAGACGCGACG	668	AGAAGCGGAGACGCGACG	668	CGUCGCUUCCGCCUUCU	1406
AAGAAAAAUAAAAGCAUUA	669	AAGAAAAAUAAAAGCAUUA	669	AUGCUUUUUUUUUUUUCUU	1407
AAUGGAGAAAAUUAAGUAGA	670	AAUGGAGAAAAUUAAGUAGA	670	UCUACUAAUUUUUCCCAUU	1408
GCUGAACAUUUUAAGACAG	671	GCUGAACAUUUUAAGACAG	671	CUGCUUUAAGAUGUUCAGC	1409
AAAAAGAAAAAUUCAGUAA	672	AAAAAGAAAAAUUCAGUAA	672	UUACUGAUUUUUUUUUUUU	1410
GAACAAGCCCCAGAAAGCC	673	GAACAAGCCCCAGAAAGCC	673	GGUCUUCUGGGGCUUUGUUC	1411
GUGAUAAAUGUCAGCUAAA	674	GUGAUAAAUGUCAGCUAAA	674	UUUAGCUGACAUUUUAUCAC	1412
GAGCCUGGAAGCAUCCAG	675	GAGCCUGGAAGCAUCCAG	675	CUGGAUGCUUCCAGGGCUC	1413
AGUGGGGGACAUCAAGCA	676	AGUGGGGGACAUCAAGCA	676	UGCUUUGAUGUCCCCCCACU	1414
GCCUGGGAGCUCUCUGGCU	677	GCCUGGGAGCUCUCUGGCU	677	AGCCAGAGAGCUCUCCAGGC	1415
UGGAAAGGACCAGCAAAGC	678	UGGAAAGGACCAGCAAAGC	678	GCUUUGCUGGUGCUUUUCCA	1416



(400/099)

AGCAGGACAUAAACAAGGUA	679	AGCAGGACAUAAACAAGGUA	679	UACCUUGUUUAUGUCCUGCU	1417
CCUAGAACUUAUAAUGCAU	680	CCUAGAACUUAUAAUGCAU	680	AUGCAUUUUAAAAGUUCUAGG	1418
AGUAGAUAAAUUAGUCAGU	681	AGUAGAUAAAUUAGUCAGU	681	ACUGACUAAUUUAUCUACU	1419
AAAUAAAGCCAGGAUUGG	682	AAAUAAAGCCAGGAUUGG	682	CCAUUCCUGGCUUUAAUUU	1420
AGUAAAUUAAAGCCAGGA	683	AGUAAAUUAAAGCCAGGA	683	UCCUGGCUUUAAUUUACU	1421
UGUGAUAAUUGUCAGCUAA	684	UGUGAUAAUUGUCAGCUAA	684	UUAGGUGACAUUUUAUCACA	1422
AGCCUGGAAGCAUCCAGG	685	AGCCUGGAAGCAUCCAGG	685	CCUGGAGCUUCCAGGGCU	1423
CACUGCUUAAAGCCUCAAUA	686	CACUGCUUAAAGCCUCAAUA	686	UAUUGAGGCUUAAAGCAGUG	1424
AAAAAUCAGUAACAGUAC	687	AAAAAUCAGUAACAGUAC	687	GUACUGUUACUGAUUUUUU	1425
GAGCCUGGAGCUCUCUGG	688	GAGCCUGGAGCUCUCUGG	688	CCAGAGAGCUCCAGGCUC	1426
UUCGCGUGGGACUUUCCA	689	UUCGCGUGGGACUUUCCA	689	UGGAAAGUCCCGAGCGGAA	1427
GAGAGACAGGCUAAUUUU	690	GAGAGACAGGCUAAUUUU	690	AAAAAUUAGCCUGUCUCUC	1428
GCUGUGAUAAUUGUCAGCU	691	GCUGUGAUAAUUGUCAGCU	691	AGCUGACAUUUUAUCACAGC	1429
CCACAGACCCCAACCCACA	692	CCACAGACCCCAACCCACA	692	UGUGGCUUUGGGGUCUCUGG	1430
CAGGAAGAAGCGGAGACAG	693	CAGGAAGAAGCGGAGACAG	693	CUGUCUCCGCUUCUUCUG	1431
UAAGCCUCAUAAAGCUUG	694	UAAGCCUCAUAAAGCUUG	694	CAAGCUUUUUAUGAGGCUUA	1432
UAAAAAGAAAAAUAUCAGU	695	UAAAAAGAAAAAUAUCAGU	695	ACUGAUUUUUUCUUUUUA	1433
GACAGAAGAAAAAUAUAAA	696	GACAGAAGAAAAAUAUAAA	696	UUUUUUUUUUUCUUCUGUC	1434
GUACCAGUAAAUUAAAGC	697	GUACCAGUAAAUUAAAGC	697	GCUUUAAUUUUACUGGUAC	1435
AAAAAGAAAAAUCAGUAAC	698	AAAAAGAAAAAUCAGUAAC	698	GUUACUGAUUUUUUCUUUUU	1436
AAAAUCAGUAACAGUACU	699	AAAAUCAGUAACAGUACU	699	AGUACUGUUUACUGAUUUUU	1437
AGAGCCUUGGAAGCAUCCA	700	AGAGCCUUGGAAGCAUCCA	700	UGGAUGCUUCCAGGGCUCU	1438
CAGGGCAAUUGGUACAUC	701	CAGGGCAAUUGGUACAUC	701	GAUGUACCAUUUUGCCCCUG	1439
CUGCAUUUACCAUACCUAG	702	CUGCAUUUACCAUACCUAG	702	CUAGGUAUGGUAUAAUGCAG	1440
UAAUUGCAUGGGUAAAAGU	703	UAAUUGCAUGGGUAAAAGU	703	ACUUUUACCCCAUGCAUUUA	1441
AAGUAAACAUAGUAACAGA	704	AAGUAAACAUAGUAACAGA	704	UCUGUUACUUAUGUUUACUU	1442
CCACACAUGCCUGUGUACC	705	CCACACAUGCCUGUGUACC	705	GGUACACAGGCAUGUGUGG	1443
AGUAGAUUUCAGAGAACUU	706	AGUAGAUUUCAGAGAACUU	706	AAGUUCUCUGAAAUCUACU	1444
CAUCAGAAAGAACCCUCCAU	707	CAUCAGAAAGAACCCUCCAU	707	AUGGAGGUUCUUUCUGAUG	1445
ACCAGUAAAAUUAAGCCA	708	ACCAGUAAAAUUAAGCCA	708	UGGCUUUAUUUUUACUGGU	1446
CACAGACCCCAACCCACAA	709	CACAGACCCCAACCCACAA	709	UUGUGGGUUGGGGUCUGUG	1447
AGGGGGAUUUGGGGGUAC	710	AGGGGGAUUUGGGGGUAC	710	GUACCCCAUUCUCCCCCU	1448
UGCAUUUACCAUACCUAGU	711	UGCAUUUACCAUACCUAGU	711	ACUAGGUUUGGUAAAUGCA	1449
CAUUGGACAUUAUCAAUUU	712	CAUUGGACAUUAUCAAUUU	712	AAAAUUUAUUGUCCAUUG	1450
CUGAACAUUUUAGACAGC	713	CUGAACAUUUUAGACAGC	713	GCUGUCUUAAAGAUUUCAG	1451
GCCUCAUAAAGCUUGCCU	714	GCCUCAUAAAGCUUGCCU	714	AGGCAAGCUUUUAUUGAGGC	1452

(400/099).

UGUACCCACAGACCCCAAC	715	UGUACCCACAGACCCCAAC	715	GUUGGGGUCUGGGGUACA	1453
GAAGUAAACUAGUACAG	716	GAAGUAAACUAGUACAG	716	CUGUUAUAUUGUUUACUUC	1454
GUAGGACCUACACCUUGA	717	GUAGGACCUACACCUUGA	717	UGACAGGUGUAGGUCCUAC	1455
CAGUGGGGGACAUAAGC	718	CAGUGGGGGACAUAAGC	718	GCUUGAUGUCCCCCACUG	1456
ACCCACUGCUUAAGCCUCA	719	ACCCACUGCUUAAGCCUCA	719	UGAGGCUUAAGCAGUGGGU	1457
AAAAUUGGGCCUGAAAAU	720	AAAAUUGGGCCUGAAAAU	720	AUUUUACAGCCCAUUUUU	1458
UGGGGGGACAUAAGCAGC	721	UGGGGGGACAUAAGCAGC	721	GCUGCUUGAUGUCCCCCA	1459
GUACAAUUGGCAGUUAUA	722	GUACAAUUGGCAGUUAUA	722	UGAAUACUGCCAUUUGUAC	1460
AAGCUAUAGGUACAGUAUU	723	AAGCUAUAGGUACAGUAUU	723	AAUACUGUACCUAUAGCUU	1461
CAGAAGAAAAUAAAAAGC	724	CAGAAGAAAAUAAAAAGC	724	GCUUUUUUUUUUUUUCUG	1462
AAUUGCAUUGGGUAAAAAGUA	725	AAUUGCAUUGGGUAAAAAGUA	725	UACUUUUACCCAUUGCAUUU	1463
AGCCUCAUAAAGCUUGCC	726	AGCCUCAUAAAGCUUGCC	726	GGCAAGCUUUUUUAGGCGU	1464
CCACUGCUUAAGCCUCAAU	727	CCACUGCUUAAGCCUCAAU	727	AUUAGGCUUUUAGCAGUGG	1465
AAGAAGCGGAGACGCGAC	728	AAGAAGCGGAGACGCGAC	728	GUCCUGUCUCGCGCUUCUU	1466
AAUUGGAGAAUUAUAGUAG	729	AAUUGGAGAAUUAUAGUAG	729	CUACUAAUUUUUUCUCCAUUU	1467
AGCCUGGGAGCUCUCUGGC	730	AGCCUGGGAGCUCUCUGGC	730	GCCAGAGAGCUCUCCAGGCU	1468
AACAAGCCCCAGAAAGCCA	731	AACAAGCCCCAGAAAGCCA	731	UGGUCUUUCUGGGGCUUGUU	1469
UACCAGUAAAUUAAAGCC	732	UACCAGUAAAUUAAAGCC	732	GGCUUUAAUUUUUACUGGUA	1470
UUCAAAAAUUGGGCCUGAA	733	UUCAAAAAUUGGGCCUGAA	733	UUCAGGCCCAUUUUUUGAA	1471
AGAAGAAAAAUAAAAAGCA	734	AGAAGAAAAAUAAAAAGCA	734	UGCUUUUUUUUUUUUCUUCU	1472
CUGUGUACCCACAGACCCC	735	CUGUGUACCCACAGACCCC	735	GGGUCUGUGGGGUACACAG	1473
GCCUGUACUGGGUCUCUCU	736	GCCUGUACUGGGUCUCUCU	736	AGAGAGACCCAGUACAGGC	1474
CAGUAAAAUUAAGCCAGG	737	CAGUAAAAUUAAGCCAGG	737	CCUGGCUUUUUUUUACUG	1475
UACAAUUGGCAGUAUUCAU	738	UACAAUUGGCAGUAUUCAU	738	AUGAAUACUGCCAUUUUGUA	1476

The 3'-ends of the Upper sequence and the Lower sequence of the siNA construct can include an overhang sequence, for example about 1, 2, 3, or 4 nucleotides in length, preferably 2 nucleotides in length, wherein the overhanging sequence of the lower sequence is optionally complementary to a portion of the target sequence. The upper sequence is also referred to as the sense strand, whereas the lower sequence is also referred to as the antisense strand. The upper and lower sequences in the Table can further comprise a chemical modification having Formulae I-VII or any combination thereof.

Table III: HIV Synthetic Modified siRNA constructs

Target Pos	Target	Seq ID	RPI#	Aliases	Sequence	Seq ID
1399	ACCAUCAUAGAGGAGGAGCUG	36		HIV43:1399U21 siRNA sense	ACCAUCAUAGAGGAGGAGCUGTT	1483
2323	UAGAUACAGGAGCAGAGAUA	8		HIV43:2323U21 siRNA sense	UAGAUACAGGAGCAGAGAUAATT	1484
2328	ACAGGAGCAGAGAUACAG	5		HIV43:2328U21 siRNA sense	ACAGGAGCAGAGAUACAGTT	1485
4930	UUUGGAAAGGACGACAAA	1		HIV43:4930U21 siRNA sense	UUUGGAAAGGACGACAAAT	1486
5077	GUAGACAGGAUAGGGAUUA	4		HIV43:5077U21 siRNA sense	GUAGACAGGAUAGGGAUUAAT	1487
5955	CUUAGGCAUCCUUAUGGC	99		HIV43:5955U21 siRNA sense	CUUAGGCAUCCUUAUGGCTT	1488
5982	GCGGAGACAGGACGGAAGA	1477		HIV43:5982U21 siRNA sense	GCGGAGACAGGACGGAAGATT	1489
8499	GCCUGUGCCUUCUACGCUA	1478		HIV43:8499U21 siRNA sense	GCCUGUGCCUUCUACGCUATT	1490
1399	ACCAUCAUAGAGGAGGAGCUG	36		HIV43:1417L21 siRNA (1399C) antisense	CAGCUUCCUACUUAUGGUTT	1491
2323	UAGAUACAGGAGCAGAGAUA	8		HIV43:2341L21 siRNA (2323C) antisense	UCAUCUGUCCUUGUAUCUATT	1492
2328	ACAGGAGCAGAGAUACAG	5		HIV43:2346L21 siRNA (2328C) antisense	CUUAGUACUUCUCCUGUATT	1493
4930	UUUGGAAAGGACGACAAA	1		HIV43:4948L21 siRNA (4930C) antisense	UUUGCUGGUCCUUAUCCAAAT	1494
5077	GUAGACAGGAUAGGGAUUA	4		HIV43:5095L21 siRNA (5077C) antisense	UAAUCCUACUCCUGUUAATT	1495
5955	CUUAGGCAUCCUUAUGGC	99		HIV43:5973L21 siRNA (5955C) antisense	GCCAUAGGAGAUAGCCUUAATT	1496
5982	GCGGAGACAGGACGGAAGA	1477		HIV43:6000L21 siRNA (5982C) antisense	UCUUCGUCGUCUCCGCTT	1497
8499	GCCUGUGCCUUCUACGCUA	1478		HIV43:8517L21 siRNA (8499C) antisense	UAGCUGAAGAGGACACAGGCTT	1498
1399	ACCAUCAUAGAGGAGGAGCUG	36		HIV43:1399U21 siRNA stab04 sense	B AccAucAucAucGAGGAAGcUtt B	1499
2323	UAGAUACAGGAGCAGAGAUA	8		HIV43:2323U21 siRNA stab04 sense	B uAGAUAcAGGAGcAGAUgATT B	1500
2328	ACAGGAGCAGAGAUACAG	5		HIV43:2328U21 siRNA stab04 sense	B AcAGGAGcAGAUgAUAcAGTT B	1501
4930	UUUGGAAAGGACGACAAA	1		HIV43:4930U21 siRNA stab04 sense	B uuUGGAAAGGAcAGcAAAT B	1502
5077	GUAGACAGGAUAGGGAUUA	4		HIV43:5077U21 siRNA stab04 sense	B GuAGAcAGGAuAGGGAuuATT B	1503
5955	CUUAGGCAUCCUUAUGGC	99		HIV43:5955U21 siRNA stab04 sense	B cuuAGGcAucuccuAuGgctt B	1504
5982	GCGGAGACAGGACGGAAGA	1477		HIV43:5982U21 siRNA stab04 sense	B GcGGAGAcAGcGAcGAAGATT B	1505
8499	GCCUGUGCCUUCUACGCUA	1478		HIV43:8499U21 siRNA stab04 sense	B GccuGuGccuucAcGcuATT B	1506
1399	ACCAUCAUAGAGGAGGAGCUG	36		HIV43:1417L21 siRNA (1399C) stab05 antisense	cAGuuccucAuuGAUGGuTsT	1507
2323	UAGAUACAGGAGCAGAGAUA	8		HIV43:2341L21 siRNA (2323C) stab05 antisense	ucAucUGuccuGuAucUAtsT	1508
2328	ACAGGAGCAGAGAUACAG	5		HIV43:2346L21 siRNA (2328C) stab05 antisense	cuGuAucAucUGuccuGuTsT	1509
4930	UUUGGAAAGGACGACAAA	1		HIV43:4948L21 siRNA (4930C) stab05 antisense	uuuGcuGGuuccuuuccAAATsT	1510
5077	GUAGACAGGAUAGGGAUUA	4		HIV43:5095L21 siRNA (5077C) stab05 antisense	uAAuccucAuccuGuuUAcTsT	1511
5955	CUUAGGCAUCCUUAUGGC	99		HIV43:5973L21 siRNA (5955C) stab05 antisense	GccAuAGGAGAUgGccuAAGTsT	1512
5982	GCGGAGACAGGACGGAAGA	1477	31236	HIV43:6000L21 siRNA (5982C) stab05 antisense	uccuocGucGuGucccGcTsT	1513
8499	GCCUGUGCCUUCUACGCUA	1478	31237	HIV43:8517L21 siRNA (8499C) stab05 antisense	uAGcuGAAGAGGAcAcAGGcTsT	1514
1399	ACCAUCAUAGAGGAGGAGCUG	36		HIV43:1399U21 siRNA stab07 sense	B AccAucAucAucGAGGAAGcUtt B	1515

2323	UAGAUACAGGAGCAGAUAGA	8		HIV43:2323U21 siRNA stab07 sense	B uAGAUACAGGAGCAGAUAGATT B	1516
2328	ACAGGAGCAGAUAGAUACAG	5		HIV43:2328U21 siRNA stab07 sense	B ACAGGAGCAGAUAGAUACAGTT B	1517
4930	UUUGGAAAGGACACAGCAAA	1		HIV43:4930U21 siRNA stab07 sense	B uuUGGAAAGGACACAGCAAAATT B	1518
5077	GUAGACAGGAUGAGGAUUA	4		HIV43:5077U21 siRNA stab07 sense	B GuAGACAGGAUGAGGAUuATT B	1519
5955	CUUAGGCAUCUCCUAUUGGC	99		HIV43:5955U21 siRNA stab07 sense	B cuuAGGCAucuccuAuGGcTT B	1520
5982	GCGGAGACAGCGACGAAGA	1477		HIV43:5982U21 siRNA stab07 sense	B GcGGAGACAGcGcAGcGAAGATT B	1521
8499	GCCUGGCCUUCUACGCUA	1478		HIV43:8499U21 siRNA stab07 sense	B GccUGuGccuucucAGcuATT B	1522
1399	ACCAUCAUAGAGGAAGCUG	36		HIV43:1417L21 siRNA (1399C) stab11 antisense	cAGcuuccuAuGauGGUtsT	1523
2323	UAGAUACAGGAGCAGAUAGA	8		HIV43:2341L21 siRNA (2323C) stab11 antisense	ucAuCuGuccuGuAuCuATsT	1524
2328	ACAGGAGCAGAUAGAUACAG	5		HIV43:2346L21 siRNA (2328C) stab11 antisense	cuGuAuAuCuGuccuGuTsT	1525
4930	UUUGGAAAGGACACAGCAAA	1		HIV43:4948L21 siRNA (4930C) stab11 antisense	uuuGcuGGuccuuuccAAATsT	1526
5077	GUAGACAGGAUGAGGAUUA	4		HIV43:5095L21 siRNA (5077C) stab11 antisense	uAuAuccuAcuccuGuCuAcTsT	1527
5955	CUUAGGCAUCUCCUAUUGGC	99		HIV43:5973L21 siRNA (5955C) stab11 antisense	GccAuAGGGAGAUgccuAAGTsT	1528
5982	GCGGAGACAGCGACGAAGA	1477	31240	HIV43:6000L21 siRNA (5982C) stab11 antisense	uccuGucGuGucuccGcTsT	1529
8499	GCCUGGCCUUCUACGCUA	1478	31241	HIV43:8517L21 siRNA (8499C) stab11 antisense	uAGcuGAAGAGGcAcAGGcTsT	1530
	AGGGGAAGUGACAUAGCAGGAAC	1479	30793	HIV:1484U21 siRNA stab04 sense	B GGGAAUGAcAuAGcAGGATT B	1531
	CAGAGUAGGAAAGGAGGGAAA	1480	30794	HIV:2666U21 siRNA stab04 sense	B GAGAUGGAAAGGAGGGGATT B	1532
	CUUGGAGGAGGAUAGAGGGA	1481	30795	HIV:7633U21 siRNA stab04 sense	B uUGGAGGAGGAGAUAGAGGTT B	1533
	CUGGAAACAUAGGAGCAUAC	1482	30796	HIV:8906U21 siRNA stab04 sense	B GGAAAAAcAuGGAGcAAucTT B	1534
	AGGGGAAGUGACAUAGCAGGAAC	1479	30797	HIV:1502L21 siRNA (1484C) stab05 antisense	uccuGcuAuGucAcuucccTsT	1535
	CAGAGUAGGAAAGGAGGGAAA	1480	30798	HIV:2684L21 siRNA (2666C) stab05 antisense	ucccuuccuuuccAuCuTsT	1536
	CUUGGAGGAGGAUAGAGGGA	1481	30799	HIV:7651L21 siRNA (7633C) stab05 antisense	ccuCuAuAuruccuccuccATsT	1537
	CUGGAAACAUAGGAGCAUAC	1482	30800	HIV:8924L21 siRNA (8906C) stab05 antisense	GAAuGcuccAuGuuuuuccTsT	1538
	ACCAUCAUAGAGGAAGCUG	36	31218	HIV43:1399U21 siRNA stab09 sense	B CAGCUCCUCAUUGAUUGGUTT B	1539
	UAGAUACAGGAGCAGAUAGA	8	31219	HIV43:2323U21 siRNA stab09 sense	B UCAUCUGCUCUCCUGUAUCUATT B	1540
	ACAGGAGCAGAUAGAUACAG	5	31220	HIV43:2328U21 siRNA stab09 sense	B CUGUAUCAUCUGCUCUCCUGUTT B	1541
	UUUGGAAAGGACACAGCAAA	1	31221	HIV43:4930U21 siRNA stab09 sense	B UUUGCUGGUCCUUUCCAAATT B	1542
	GUAGACAGGAUGAGGAUUA	4	31222	HIV43:5077U21 siRNA stab09 sense	B UAAUCCUCAUCCUGUCUACTT B	1543
	CUUAGGCAUCUCCUAUUGGC	99	31223	HIV43:5955U21 siRNA stab09 sense	B GCCAUAGGAGAUAGCCUAAGTT B	1544
	GCGGAGACAGCGACGAAGA	1477	31224	HIV43:5982U21 siRNA stab09 sense	B UCUUCGUCGUCUGUCUCCGCTT B	1545
	GCCUGGCCUUCUACGCUA	1478	31225	HIV43:8499U21 siRNA stab09 sense	B UAGCUGAAGAGGACACAGGCTT B	1546
	ACCAUCAUAGAGGAAGCUG	36	31226	HIV43:1417L21 siRNA (1399C) stab10 antisense	CAGCUCCUCAUUGAUUGGUTsT	1547
	UAGAUACAGGAGCAGAUAGA	8	31227	HIV43:2341L21 siRNA (2323C) stab10 antisense	UCAUCUCCUCCUGUAUCUATsT	1548
	ACAGGAGCAGAUAGAUACAG	5	31228	HIV43:2346L21 siRNA (2328C) stab10 antisense	CUGUAUCAUCUGCUCUCCUGUTsT	1549
	UUUGGAAAGGACACAGCAAA	1	31229	HIV43:4948L21 siRNA (4930C) stab10 antisense	UUUGCUGGUCCUUUCCAAATsT	1550
	GUAGACAGGAUGAGGAUUA	4	31230	HIV43:5095L21 siRNA (5077C) stab10 antisense	UAAUCCUCAUCCUGUCUACTsT	1551
	CUUAGGCAUCUCCUAUUGGC	99	31231	HIV43:5973L21 siRNA (5955C) stab10 antisense	GCCAUAGGAGAUAGCCUAAGTsT	1552
	GCGGAGACAGCGACGAAGA	1477	31232	HIV43:6000L21 siRNA (5982C) stab10 antisense	UCUUCGUCGUCUGUCUCCGCTsT	1553

(400/099)

	GCCUGUGCCUUCUACGCUA	1478	31233	HIV43:8517L21 siRNA (8499C) stab10 antisense	UAGCUGAAGAGGACACAGGCTsT	1554
	GCGGAGACAGCGACGAAGA	1477	31234	HIV43:5982U21 siRNA stab04 sense	B ucuucGucGcuGuccGcTt B	1555
	GCCUGUGCCUUCUACGCUA	1478	31235	HIV43:8499U21 siRNA stab04 sense	B uAGCUGAAGAGGAcAcAGGcTt B	1556
	GCGGAGACAGCGACGAAGA	1477	31238	HIV43:5982U21 siRNA stab07 antisense	B ucuucGucGcuGuccGcTt B	1557
	GCCUGUGCCUUCUACGCUA	1478	31239	HIV43:8499U21 siRNA stab07 antisense	B uAGCUGAAGAGGAcAcAGGcTt B	1558
	ACCAUCAUAGAGGAGCUG	36	31242	HIV43:1399U21 siRNA inv stab09 sense	B GUCGAAAGGAGUAACUACCATt B	1559
	UAGAUACAGGAGCAGAUAGA	8	31243	HIV43:2323U21 siRNA inv stab09 sense	B AGUAGACGAGGACAUAGAUTt B	1560
	ACAGGAGCAGAGAUACAG	5	31244	HIV43:2328U21 siRNA inv stab09 sense	B GACAUAUAGACGAGGACATt B	1561
	UUUGGAAAGGACCAAGCAAA	1	31245	HIV43:4930U21 siRNA inv stab09 sense	B AAACGACCAGGAAAGGUUUTt B	1562
	GUAGACAGGAUGAGGAUUA	4	31246	HIV43:5077U21 siRNA inv stab09 sense	B AUUAGGAGUAGGACAGAUUTt B	1563
	CUUAGGCAUCUCCUAUGGC	99	31247	HIV43:5955U21 siRNA inv stab09 sense	B CGGUAUCCUACGGAUJUCTt B	1564
	GCGGAGACAGCGACGAAGA	1477	31248	HIV43:5982U21 siRNA inv stab09 sense	B AGAAGCAGCGACAGAGGCGTt B	1565
	GCCUGUGCCUUCUACGCUA	1478	31249	HIV43:8499U21 siRNA inv stab09 sense	B AUCGACUUCUCCGUGUCCGTT B	1566
	ACCAUCAUAGAGGAGCUG	36	31250	HIV43:1417L21 siRNA (1399C) inv stab10 antisense	UGGUAGUUACUCCUUCGACTsT	1567
	UAGAUACAGGAGCAGAUAGA	8	31251	HIV43:2341L21 siRNA (2323C) inv stab10 antisense	AUCUAUGUCCUCGUCUACUTsT	1568
	ACAGGAGCAGAUACAG	5	31252	HIV43:2346L21 siRNA (2328C) inv stab10 antisense	UGUCCUCGUCUACUAGUGUCTsT	1569
	UUUGGAAAGGACCAAGCAAA	1	31253	HIV43:4948L21 siRNA (4930C) inv stab10 antisense	AAACCUUUCUGGUGCGUUUTsT	1570
	GUAGACAGGAUGAGGAUUA	4	31254	HIV43:5095L21 siRNA (5077C) inv stab10 antisense	CAUCUGUCUACUCCUAAUTsT	1571
	CUUAGGCAUCUCCUAUGGC	99	31255	HIV43:5973L21 siRNA (5955C) inv stab10 antisense	GAUCCGUAAGAGGAUACCGTsT	1572
	GCGGAGACAGCGACGAAGA	1477	31256	HIV43:6000L21 siRNA (5982C) inv stab10 antisense	CGCCUCUGUCGUGCUUUTsT	1573
	GCCUGUGCCUUCUACGCUA	1478	31257	HIV43:8517L21 siRNA (8499C) inv stab10 antisense	CGGACACGGAGAAGUCGAUTsT	1574
	GCGGAGACAGCGACGAAGA	1477	31258	HIV43:5982U21 siRNA inv stab04 sense	B AGAAGcAGGAcAGAGGcGTT B	1575
	GCCUGUGCCUUCUACGCUA	1478	31259	HIV43:8499U21 siRNA inv stab04 sense	B AucGAcuucuccGuccGTT B	1576
	GCGGAGACAGCGACGAAGA	1477	31260	HIV43:6000L21 siRNA (5982C) inv stab05 antisense	cGccuucGucGcuGauucUtsT	1577
	GCCUGUGCCUUCUACGCUA	1478	31261	HIV43:8517L21 siRNA (8499C) inv stab05 antisense	cGGAcAcGGAGAAAGucGAUTsT	1578
	GCGGAGACAGCGACGAAGA	1477	31262	HIV43:5982U21 siRNA inv stab07 sense	B AGAAGcAGGAcAGAGGcGTT B	1579
	GCCUGUGCCUUCUACGCUA	1478	31263	HIV43:8499U21 siRNA inv stab07 sense	B AucGAcuucuccGuccGTT B	1580
	GCGGAGACAGCGACGAAGA	1477	31264	HIV43:6000L21 siRNA (5982C) inv stab11 antisense	cGccuucGucGcuGauucUtsT	1581
	GCCUGUGCCUUCUACGCUA	1478	31265	HIV43:8517L21 siRNA (8499C) inv stab11 antisense	cGGAcAcGGAGAAAGucGAUTsT	1582

Uppercase = ribonucleotide

u,c = 2'-deoxy-2'-fluoro U, C

T = thymidine

B = inverted deoxy abasic

s = phosphorothioate linkage

A = deoxy Adenosine

G = deoxy Guanosine

(400/099)

**Table IV**

Non-limiting examples of Stabilization Chemistries for chemically modified siNA constructs

Chemistry	Pyrimidine	Purine	Cap	p=S	Strand
"Stab 1"	Ribo	Ribo	-	5 at 5'-end 1 at 3'-end	S/AS
"Stab 2"	Ribo	Ribo	-	All linkages	Usually AS
"Stab 3"	2'-fluoro	Ribo	-	4 at 5'-end 4 at 3'-end	Usually S
"Stab 4"	2'-fluoro	Ribo	5' and 3'-ends	-	Usually S
"Stab 5"	2'-fluoro	Ribo	-	1 at 3'-end	Usually AS
"Stab 6"	2'-O-Methyl	Ribo	5' and 3'-ends	-	Usually S
"Stab 7"	2'-fluoro	2'-deoxy	5' and 3'-ends	-	Usually S
"Stab 8"	2'-fluoro	2'-O-Methyl	-	1 at 3'-end	Usually AS
"Stab 9"	Ribo	Ribo	5' and 3'-ends	-	Usually S
"Stab 10"	Ribo	Ribo	-	1 at 3'-end	Usually AS
"Stab 11"	2'-fluoro	2'-deoxy	-	1 at 3'-end	Usually AS

CAP = any terminal cap, see for example **Figure 10**.

All Stab 1-11 chemistries can comprise 3'-terminal thymidine (TT) residues

All Stab 1-11 chemistries typically comprise 21 nucleotides, but can vary as described herein.

S = sense strand

AS = antisense strand

**Table V**A. 2.5  $\mu$ mol Synthesis Cycle ABI 394 Instrument

Reagent	Equivalents	Amount	Wait Time* DNA	Wait Time* 2'-O-methyl	Wait Time*RNA
Phosphoramidites	6.5	163 $\mu$ L	45 sec	2.5 min	7.5 min
S-Ethyl Tetrazole	23.8	238 $\mu$ L	45 sec	2.5 min	7.5 min
Acetic Anhydride	100	233 $\mu$ L	5 sec	5 sec	5 sec
N-Methyl Imidazole	186	233 $\mu$ L	5 sec	5 sec	5 sec
TCA	176	2.3 mL	21 sec	21 sec	21 sec
Iodine	11.2	1.7 mL	45 sec	45 sec	45 sec
Beaucage	12.9	645 $\mu$ L	100 sec	300 sec	300 sec
Acetonitrile	NA	6.67 mL	NA	NA	NA

B. 0.2  $\mu$ mol Synthesis Cycle ABI 394 Instrument

Reagent	Equivalents	Amount	Wait Time* DNA	Wait Time* 2'-O-methyl	Wait Time*RNA
Phosphoramidites	15	31 $\mu$ L	45 sec	233 sec	465 sec
S-Ethyl Tetrazole	38.7	31 $\mu$ L	45 sec	233 min	465 sec
Acetic Anhydride	655	124 $\mu$ L	5 sec	5 sec	5 sec
N-Methyl Imidazole	1245	124 $\mu$ L	5 sec	5 sec	5 sec
TCA	700	732 $\mu$ L	10 sec	10 sec	10 sec
Iodine	20.6	244 $\mu$ L	15 sec	15 sec	15 sec
Beaucage	7.7	232 $\mu$ L	100 sec	300 sec	300 sec
Acetonitrile	NA	2.64 mL	NA	NA	NA

C. 0.2  $\mu$ mol Synthesis Cycle 96 well Instrument

Reagent	Equivalents:DNA/ 2'-O-methyl/Ribo	Amount: DNA/2'-O- methyl/Ribo	Wait Time* DNA	Wait Time* 2'-O- methyl	Wait Time* Ribo
Phosphoramidites	22/33/66	40/60/120 $\mu$ L	60 sec	180 sec	360sec
S-Ethyl Tetrazole	70/105/210	40/60/120 $\mu$ L	60 sec	180 min	360 sec
Acetic Anhydride	265/265/265	50/50/50 $\mu$ L	10 sec	10 sec	10 sec
N-Methyl Imidazole	502/502/502	50/50/50 $\mu$ L	10 sec	10 sec	10 sec
TCA	238/475/475	250/500/500 $\mu$ L	15 sec	15 sec	15 sec
Iodine	6.8/6.8/6.8	80/80/80 $\mu$ L	30 sec	30 sec	30 sec
Beaucage	34/51/51	80/120/120	100 sec	200 sec	200 sec
Acetonitrile	NA	1150/1150/1150 $\mu$ L	NA	NA	NA

- 5
- Wait time does not include contact time during delivery.
  - Tandem synthesis utilizes double coupling of linker molecule

CLAIMS

What we claim is:

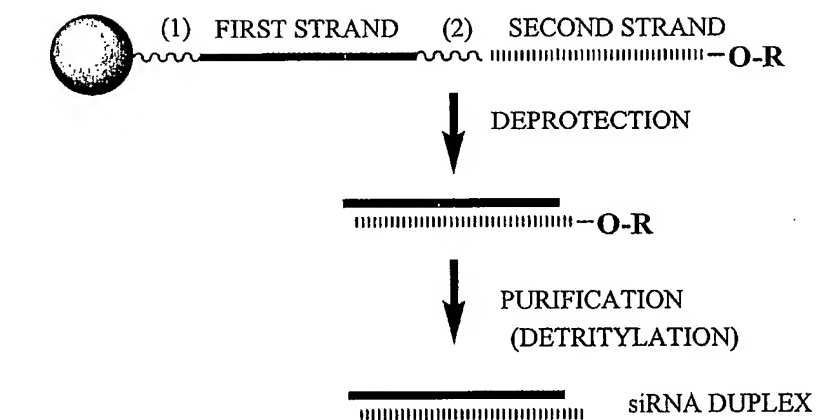
1. A double-stranded short interfering nucleic acid (siNA) molecule that inhibits replication of a human immunodeficiency virus (HIV), wherein one of the strands of said double-stranded siNA molecule is an antisense strand which comprises a nucleotide sequence that is complementary to the nucleotide sequence of an HIV RNA or a portion thereof and the other strand is a sense strand which comprises a nucleotide sequence that is complementary to the nucleotide sequence of the antisense strand and wherein a majority of the pyrimidine nucleotides present in said double-stranded siNA molecule comprises a sugar modification.
2. The siNA molecule of claim 1, wherein the HIV RNA comprises HIV-1 RNA.
3. The siNA molecule of claim 1, wherein the HIV RNA comprises HIV-2 RNA.
4. The siNA molecule of claim 1, wherein the siNA molecule comprises no ribonucleotides.
5. The siNA molecule of claim 1, wherein the siNA molecule comprises ribonucleotides.
6. The siNA molecule of claim 1, wherein all the pyrimidine nucleotides in the siNA molecule comprise sugar modifications.
7. The siNA molecule of claim 6, wherein the modified pyrimidine nucleotides are selected from 2'-deoxy-pyrimidine, 2'-O-alkyl pyrimidine, 2'-C-alkyl pyrimidine, 2'-deoxy-2'-fluoro-pyrimidine, 2'-amino pyrimidine, 2'-methoxy-ethoxy pyrimidine, and 2'-O, 4'-C-methylene pyrimidine nucleotides, alone or in any combination thereof.
8. The siNA molecule of claim 7, wherein the 2'-O-alkyl pyrimidine nucleotide is 2'-O-methyl or 2'-O-allyl.
9. The siNA molecule of claim 1, wherein the nucleotide sequence of the antisense strand of the double-stranded siNA molecule is complementary to an RNA encoding an HIV protein or a fragment thereof.
10. The siNA molecule of claim 1, wherein each strand of the siNA molecule comprises about 19 to about 29 nucleotides, and wherein each strand comprises at least about 19 nucleotides that are complementary to the nucleotides of the other strand.



11. The siNA molecule of claim 1, wherein said siNA molecule is assembled from two oligonucleotide fragments, wherein one fragment comprises the nucleotide sequence of the antisense strand of the siNA molecule and the second fragment comprises the nucleotide sequence of the sense strand of the siNA molecule.
- 5 12. The siNA molecule of claim 1, wherein the sense strand is connected to the antisense strand via a linker molecule.
13. The siNA molecule of claim 12, wherein said linker molecule is a polynucleotide linker.
14. The siNA molecule of claim 12, wherein said linker molecule is a non-nucleotide  
10 linker.
15. The siNA molecule of claim 1, wherein any pyrimidine nucleotides present in the sense strand are 2'-deoxy-2'-fluoro pyrimidine nucleotides and wherein any purine nucleotides present in the sense region are 2'-deoxy purine nucleotides.
16. The siNA molecule of claim 1, wherein the sense strand comprises a 3'-end and a 5'-  
15 end, and wherein a terminal cap moiety is present at the 5'-end, the 3'-end, or both of the 5' and 3' ends of said sense strand.
17. The siNA molecule of claim 16, wherein said terminal cap moiety is an inverted deoxy abasic moiety.
18. The siNA molecule of claim 1, wherein the antisense strand comprises one or more  
20 2'-deoxy-2'-fluoro pyrimidine nucleotides and one or more 2'-O-methyl purine nucleotides.
19. The siNA molecule of claim 1, wherein any pyrimidine nucleotides present in the antisense strand are 2'-deoxy-2'-fluoro pyrimidine nucleotides and wherein any purine nucleotides present in the antisense strand are 2'-O-methyl purine nucleotides.
- 25 20. The siNA molecule of claim 1, wherein the antisense strand comprises a phosphorothioate internucleotide linkage at the 3' end of said antisense strand.
21. The siNA molecule of claim 1, wherein the antisense strand comprises a glyceryl modification at the 3' end of said antisense strand.
22. The siNA molecule of claim 1, wherein each strand of the siNA molecule comprises  
30 21 nucleotides.

23. The siNA molecule of claim 22, wherein about 19 nucleotides of each strand of the siNA molecule are base-paired to the complementary nucleotides of the other strand of the siNA molecule and wherein at least two 3' terminal nucleotides of each strand of the siNA molecule are not base-paired to the nucleotides of the other strand of the siNA molecule.
24. The siNA molecule of claim 23, wherein each of the two 3' terminal nucleotides of each fragment of the siNA molecule are 2'-deoxy-pyrimidines.
25. The siNA molecule of claim 24, wherein the 2'-deoxy-pyrimidine is 2'-deoxy-thymidine.
26. The siNA molecule of claim 22, wherein 21 nucleotides of each strand of the siNA molecule are base-paired to the complementary nucleotides of the other strand of the siNA molecule.
27. The siNA molecule of claim 22, wherein about 19 nucleotides of the antisense strand are base-paired to the nucleotide sequence of an HIV RNA or a portion thereof.
28. The siNA molecule of claim 22, wherein 21 nucleotides of the antisense strand are base-paired to the nucleotide sequence of an HIV RNA or a portion thereof.
29. The siNA molecule of claim 1, wherein the 5'-end of the antisense strand optionally includes a phosphate group.
30. The siNA molecule of claim 1, wherein the nucleotide sequence of the antisense strand or a portion thereof is complementary to the nucleotide sequence of the 5'-untranslated region of an HIV RNA or a portion thereof.
31. The siNA molecule of claim 1, wherein the nucleotide sequence of the antisense strand or a portion thereof is complementary to the nucleotide sequence of an HIV RNA or a portion thereof that is present in the RNA of all HIV.
32. A pharmaceutical composition comprising the siNA molecule of claim 1 in an acceptable carrier or diluent.
33. A medicament comprising the siNA molecule of claim 1.
34. An active ingredient comprising the siNA molecule of claim 1.
35. The use of a double-stranded short interfering nucleic acid (siNA) molecule that inhibits replication of a human immunodeficiency virus (HIV), wherein one of the

5 strands of the double-stranded siNA molecule is an antisense strand which comprises a nucleotide sequence that is complementary to the nucleotide sequence of an HIV RNA or a portion thereof and the other strand is a sense strand which comprises a nucleotide sequence that is complementary to the nucleotide sequence of the antisense strand and wherein a majority of the pyrimidine nucleotides present in the double-stranded siNA molecule comprises a sugar modification.

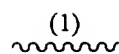
*Figure 1*

= SOLID SUPPORT

R = TERMINAL PROTECTING GROUP

FOR EXAMPLE:

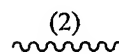
DIMETHOXYTRITYL (DMT)



= CLEAVABLE LINKER

(FOR EXAMPLE: NUCLEOTIDE SUCCINATE OR

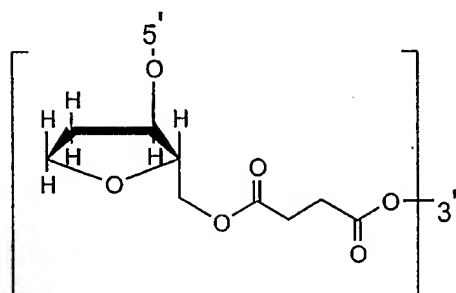
INVERTED DEOXYABASIC SUCCINATE)



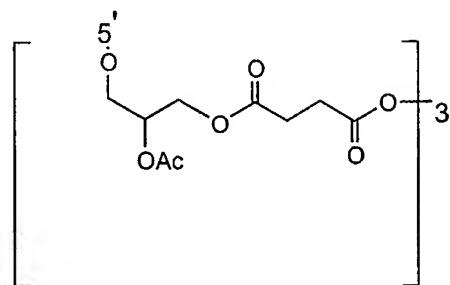
= CLEAVABLE LINKER

(FOR EXAMPLE: NUCLEOTIDE SUCCINATE OR

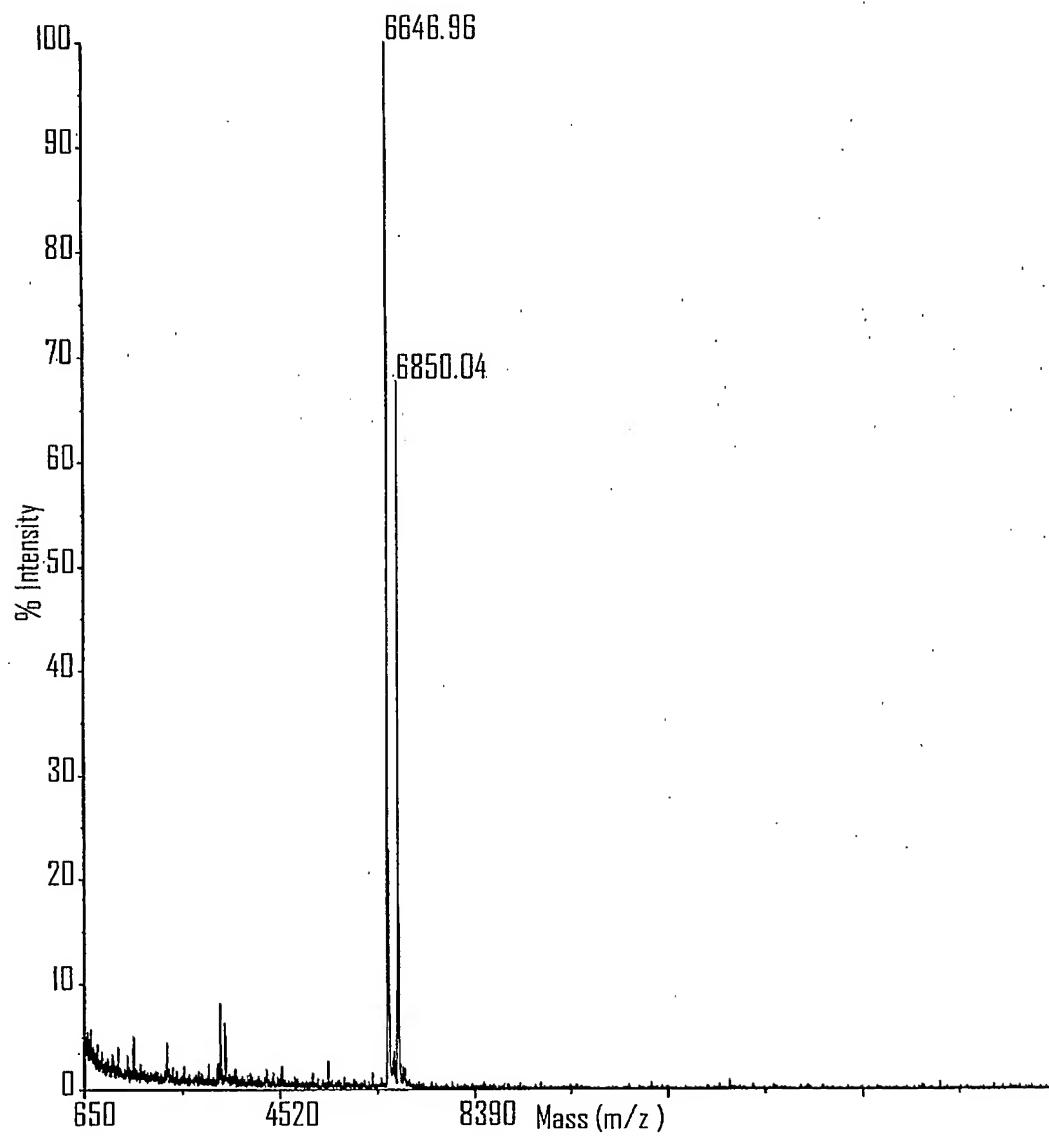
INVERTED DEOXYABASIC SUCCINATE)



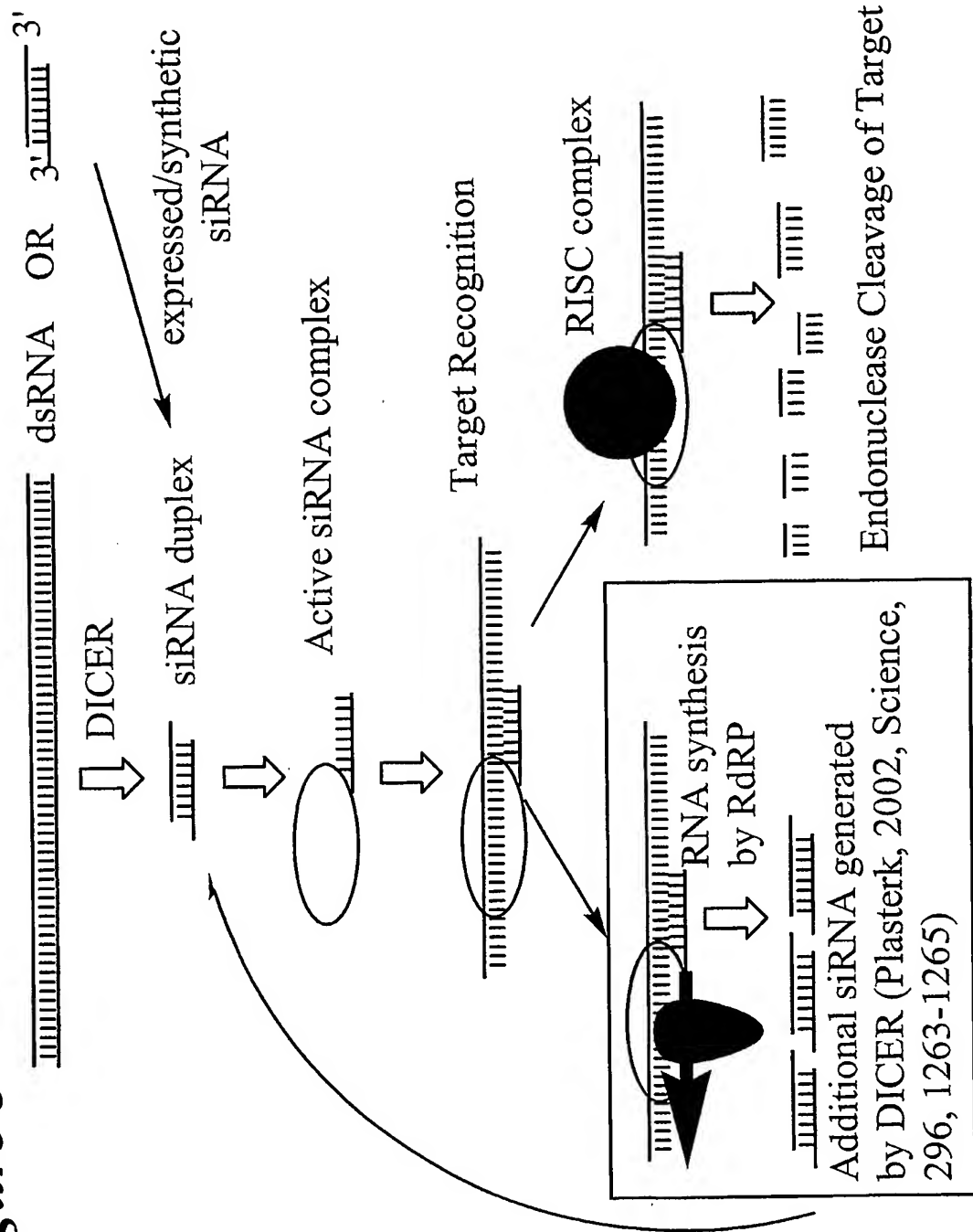
INVERTED DEOXYABASIC SUCCINATE  
LINKAGE



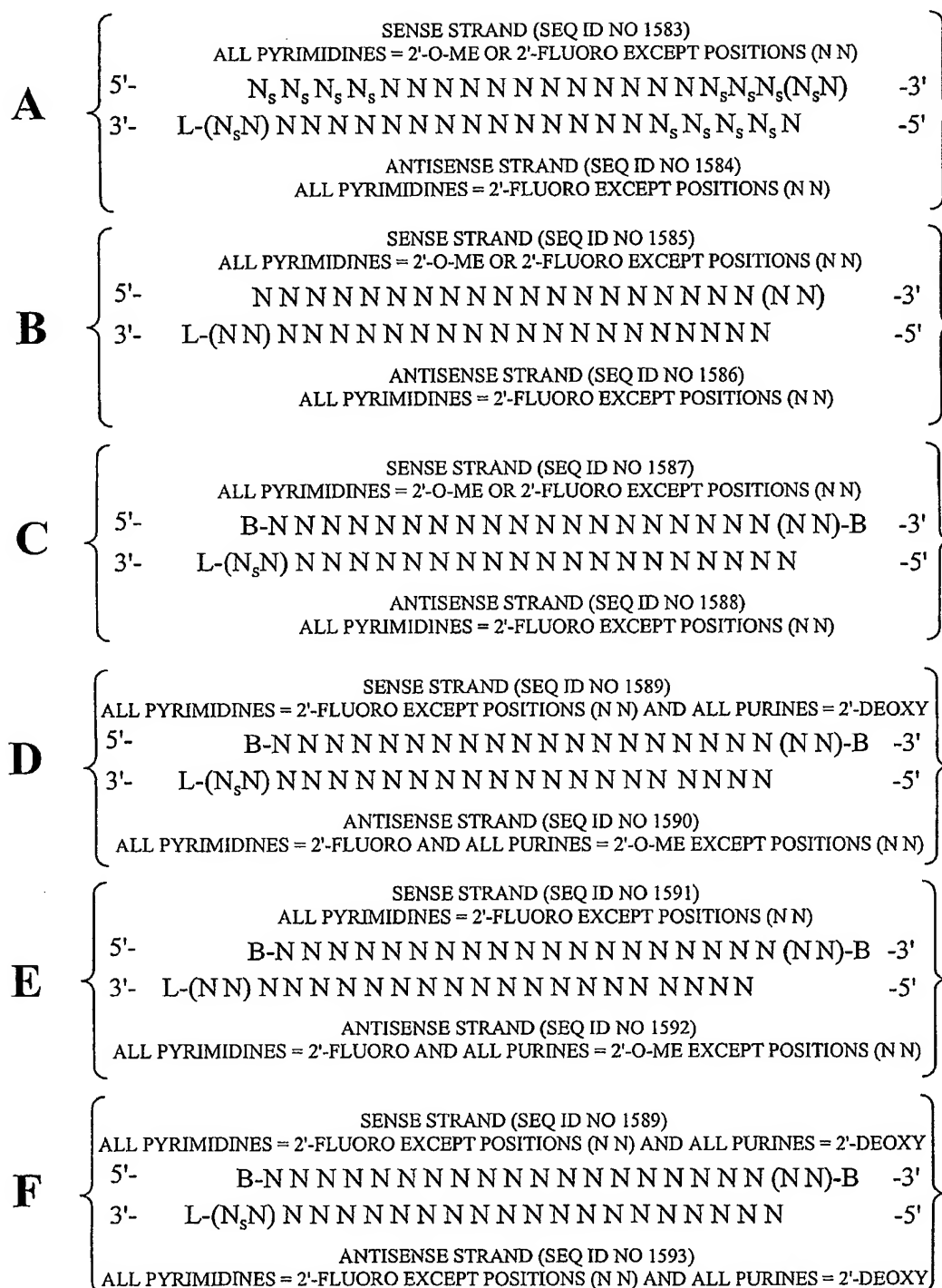
GLYCERYL SUCCINATE LINKAGE

*Figure 2*

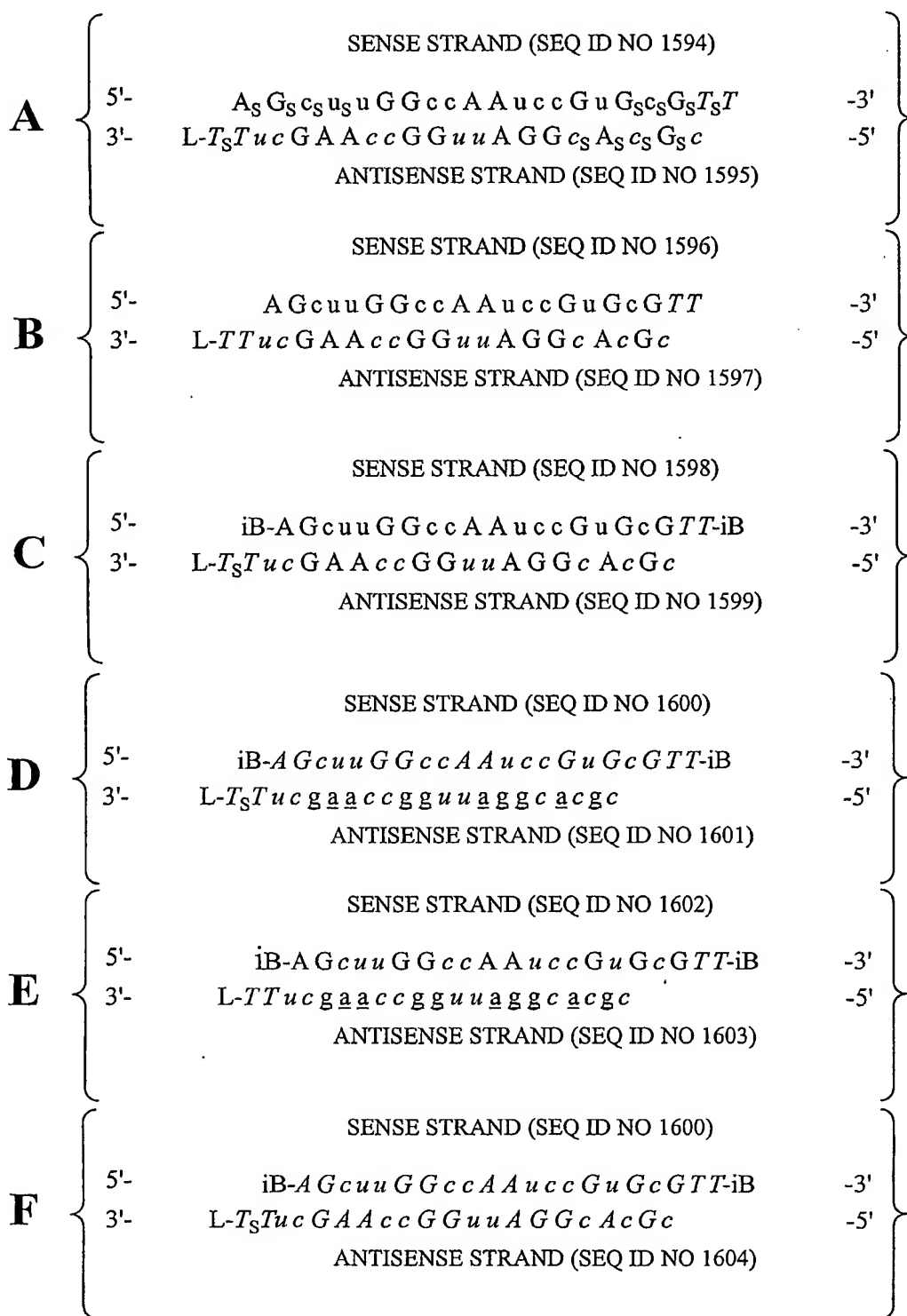
**Figure 3**



**Figure 4**



POSITIONS (NN) CAN COMPRISE ANY NUCLEOTIDE, SUCH AS DEOXYNUCLEOTIDES  
(eg. THYMIDINE) OR UNIVERSAL BASES  
B = ABASIC, INVERTED ABASIC, INVERTED NUCLEOTIDE OR OTHER TERMINAL CAP  
THAT IS OPTIONALLY PRESENT  
L = GLYCERYL MOIETY THAT IS OPTIONALLY PRESENT  
S = PHOSPHOROTHIOATE OR PHOSPHORODITHIOATE

**Figure 5**

lower case = 2'-O-Methyl or 2'-deoxy-2'-fluoro

*italic lower case* = 2'-deoxy-2'-fluorounderline = 2'-O-methyl*ITALIC UPPER CASE* = DEOXY

B = INVERTED DEOXYABASIC

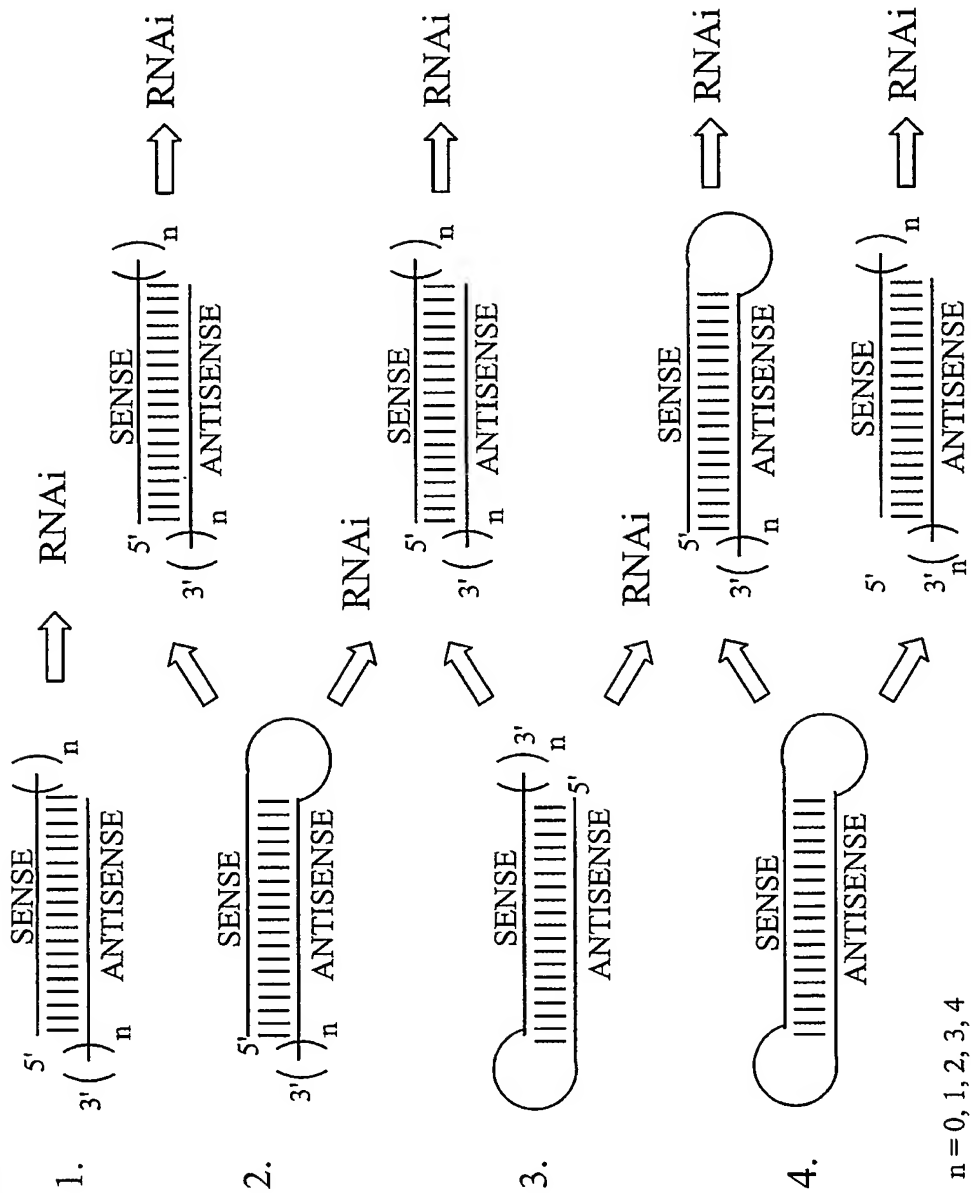
L = GLYCERYL MOIETY OPTIONALLY PRESENT

S = PHOSPHOROTHIOATE OR

PHOSPHORODITHIOATE

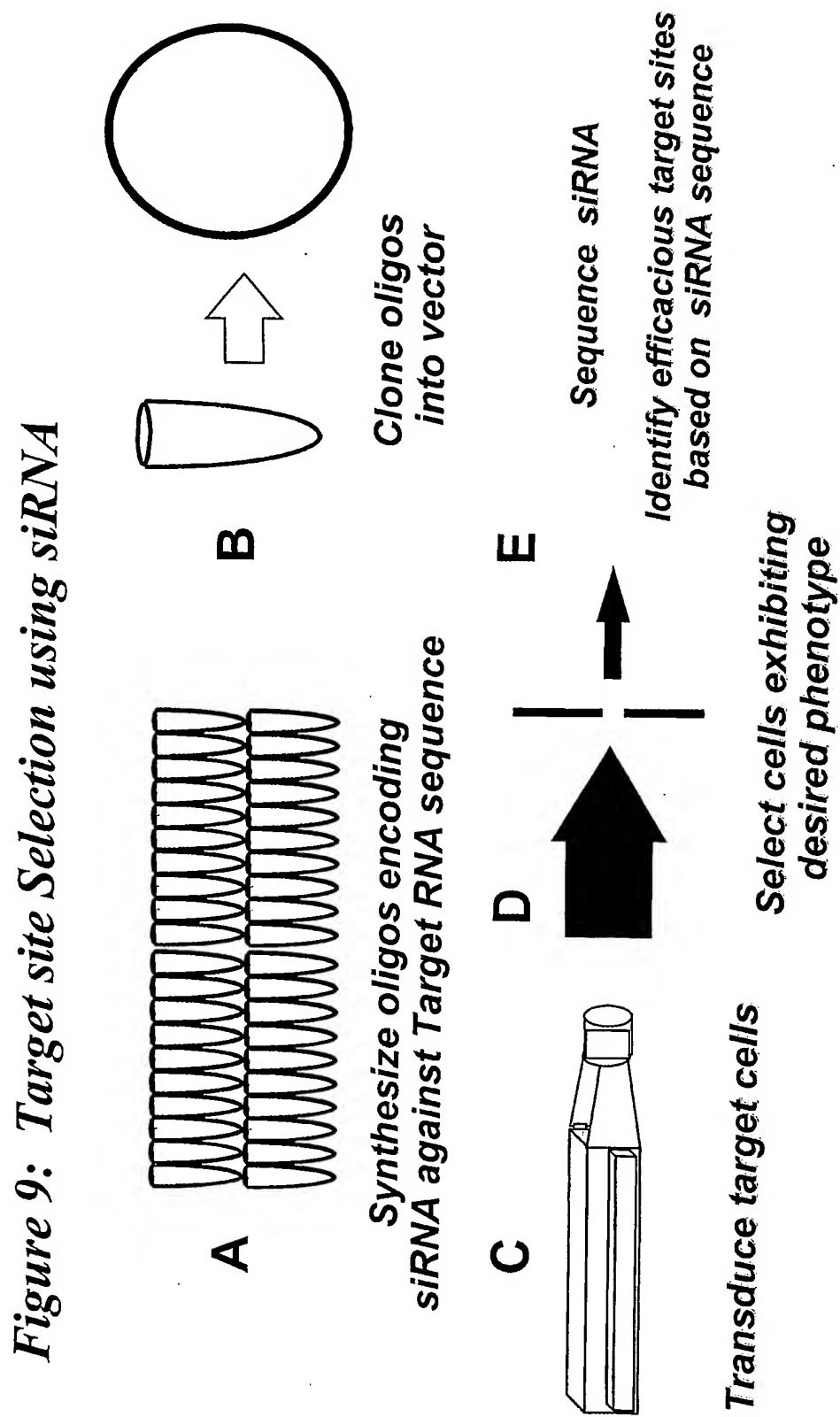


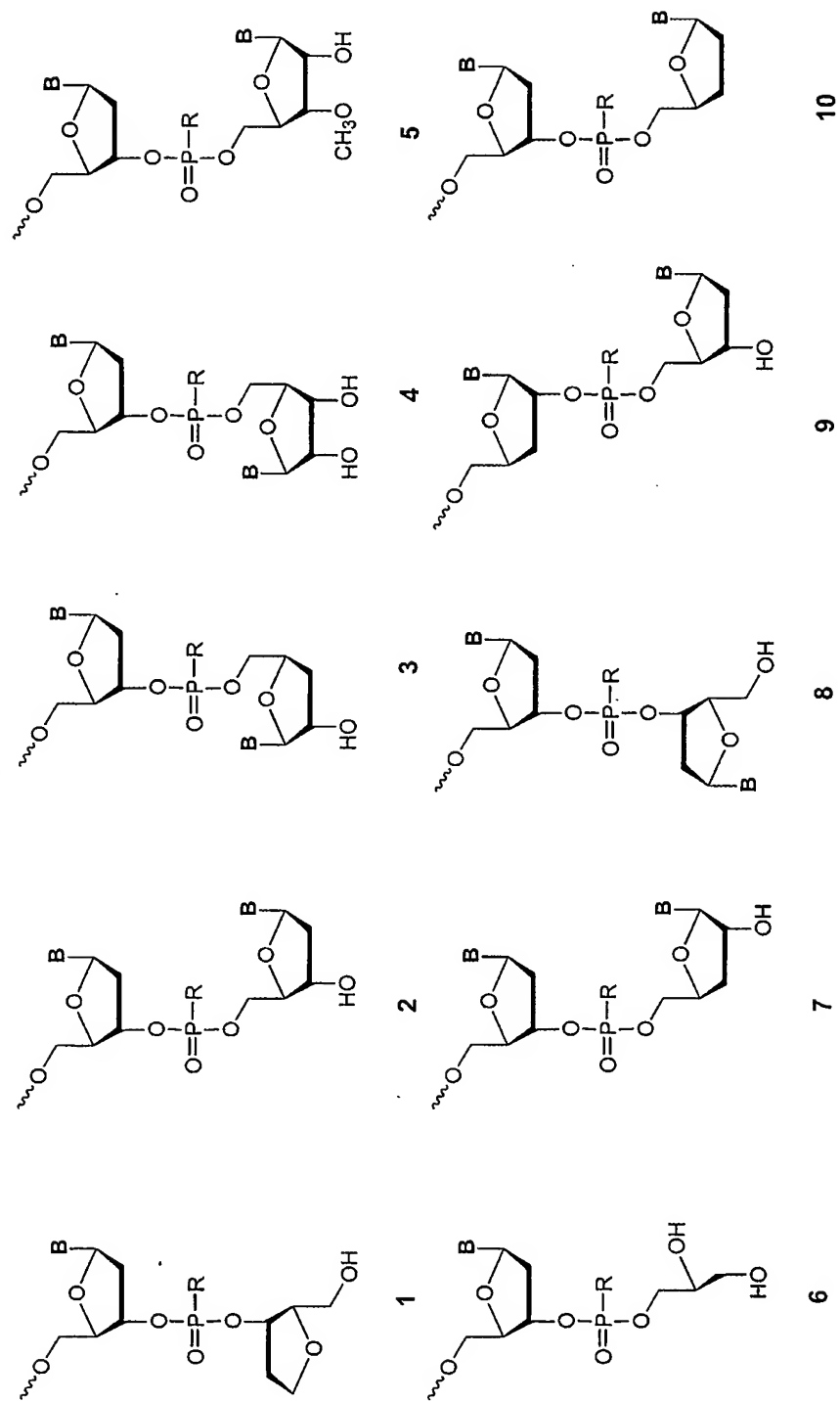
**Figure 6**









*Figure 10*

R = O, S, N, alkyl, substituted alkyl, O-alkyl, S-alkyl, alkaryl, or aralkyl  
B = Independently any nucleotide base, either naturally occurring or chemically modified, or optionally H (abasic).

*Figure 11: Modification Strategy*

